(19)日本国特許庁（ＪＰ）　(12)公開特許公報（Ａ）　(11)特許出願公開番号

特開平9－200239

(43)公開日　平成９年(1997)７月31日

(51)Int.Cl.[6]		識別記号	庁内整理番号	ＦＩ			技術表示箇所
Ｈ０４Ｌ	12/42			Ｈ０４Ｌ	11/00	３３０	
Ｇ０６Ｆ	13/00	３５７		Ｇ０６Ｆ	13/00	３５７Ｃ	
	13/36	５３０			13/36	５３０Ｃ	

審査請求　未請求　請求項の数９　ＯＬ　（全 44 頁）

(21)出願番号　特願平8－7140

(22)出願日　平成８年(1996)１月19日

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地

(72)発明者　岡田　康行
神奈川県海老名市下今泉810番地株式会社日立製作所オフィスシステム事業部内

(74)代理人　弁理士　小川　勝男

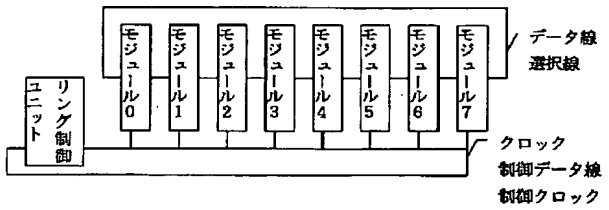
(54)【発明の名称】　リング接続を用いたデータ転送方法及び情報処理システム

(57)【要約】

【課題】長距離のプロセッサ間の接続方法であったリング接続をプロセッサ間の高速接続に適用する際の隘路となっていたレイテンシ、スループット、運用機能を改善し、プロセッサ間を高速に接続する。

【解決手段】複数の信号線からなるリングでプロセッサ等のモジュールを複数接続し、送信権を獲得するためのフラグ用の独立する信号線を設け、前記フラグをセットして送信権を要求し、受信したフラグがセットされていないことから前記送信権を先行して獲得したことを後から確認することで受信及び送信動作と前記リングの転送動作を並行して行う。

図２４



1

【特許請求の範囲】

【請求項１】複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおけるリング接続を用いたデータ転送方法であって、送信権を獲得するためのフラグ用の独立する信号線を設け、前記フラグをセットして送信権を要求し、受信したフラグがセットされていないことから前記送信権を先行して獲得したことを後から確認することで受信及び送信動作と前記リングの転送動作を並行して行うことを特徴とするリング接続を用いたデータ転送方法。

【請求項２】複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおけるリング接続を用いたデータ転送方法であって、同一クロックを前記リングに配信し、転送及び送受信に使用することを特徴とするリング接続を用いたデータ転送方法。

【請求項３】複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおけるリング接続を用いたデータ転送方法であって、位相調整回路によりクロックの位相を受信する信号に合わせ、前記クロックの乗り換えの遅延を不要にするためクロックを転送制御、受信制御、送信制御に同一位相で使用し、前記複数のモジュールのうちの少なくとも一つのモジュールは回送モードで前記リング１周の遅延を前記クロックの整数倍のサイクルに合わせることを特徴とするリング接続を用いたデータ転送方法。

【請求項４】複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおけるリング接続を用いたデータ転送方法であって、前記複数のモジュールのうちの各モジュールはラッチせずに前記複数のモジュールのうちの次のモジュールに転送し、ラッチ前の位相に合わせて送信することを特徴とするリング接続を用いたデータ転送方法。

【請求項５】複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおけるリング接続を用いた情報処理システムであって、送信権を獲得するためのフラグ用の独立する信号線を設け、前記フラグをセットして送信権を要求し、受信したフラグがセットされていないことから前記送信権を先行して獲得したことを後から確認する手段を有することを特徴とするリング接続を用いた情報処理システム。

【請求項６】複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおけるリング接続を用いた情報処理システムであって、クロックの位相を受信する信号に合わせる位相調整回路と、転送制御、受信制御、送信制御に同一位相で使用するクロックを供給する手段と、前記リング１周の遅延を前記クロックの整数倍のサイクルに合わせる少なくとも一つの回送モードのモジュールを有することを特徴とするリング接続を用いた情報処理システム。

2

【請求項７】複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおけるリング接続を用いた情報処理システムであって、前記複数のモジュールのうちの各モジュールは、ラッチせずに前記複数のモジュールのうちの次のモジュールに転送する手段と、ラッチ前の位相に合わせて送信する手段とを有することを特徴とするリング接続を用いた情報処理システム。

【請求項８】複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおけるリング接続を用いた情報処理システムであって、前記複数のモジュールのうちの各モジュールは、ラッチせずに前記複数のモジュールのうちの次のモジュールに転送する手段と、ラッチ前の位相に合わせて送信する手段とを有することを特徴とするリング接続を用いた情報処理システム。

【請求項９】複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおけるリング接続を用いた情報処理システムであって、パケットを連続的に転送するブロック・パケットの送信権を獲得するための信号線と、前記送信権を獲得した２サイクル後からブロック転送の間通常のパケットの送信要求を抑止させることを通知する信号線と、ブロックパケットの送信権の獲得後の４サイクル後からブロックパケットの送信を許可する手段とを有することを特徴とするリング接続を用いた情報処理システム。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、パーソナルコンピュータ、ワークステーション、サーバ等の情報処理装置に用いられるデータ転送方法及びそれを用いた情報処理システムに関し、特に複数のプロセッサ間のデータ転送を高速に行うのに好適なリング接続を用いたデータ転送方法及びそれを用いた情報処理システムに関する。

【０００２】

【従来の技術】パーソナルコンピュータ等の情報処理装置が普及し、様々な分野で使用されるようになるにつれて、情報処理装置の処理性能の向上が要求されている。性能向上の手段として、情報処理装置を複数のプロセッサから構成するアプローチがある。この構成の場合、複数のプロセッサ間のデータ転送を高速に行うことが性能上の重要な鍵となる。プロセッサ間の高速接続手段として、リング接続方式と現在使用されている接続方式であるバス、スイッチとの比較を図１に示す。

【０００３】プロセッサ間データ転送を高速に行うためのプロセッサ間の接続手段として、バスが最も広く用いられている。バスは外部に付加回路を使用せず接続でき、拡張性に富むとともに、バス制御ＬＳＩ、コネクタなどに必要なピンの数が少なく低価格である。また、すべてのプロセッサは、バスの内容を見ることができるた

め密結合のプロセッサ間接続の全報知の通信（スヌープ）に適する。

【０００４】バス接続の場合、性能、接続台数を決める主要な要因である通信容量は、動作周波数に依存する。プロセッサバスでは直近のプロセッサと最遠のプロセッサからの信号伝搬の時間差を同一クロック内に納める必要があり、動作周波数にはバスの長さによる限界がある。クロック、ＬＳＩ、伝送線のスキューの設計値は、動作周波数に直接影響を与える。送信元と受信先にすべての組み合わせがあるため、スキュー補正に適当な手段がない。プロセッサへの引出線からの信号の両方向の伝搬、引出線の通過により伝搬特性の劣化、すべてのプロセッサへの駆動能力が必要なことによる動作周波数の限界もある。バスの長さを短くして動作周波数を向上させることも行われているが、マイクロプロセッサ・システムの実装形態に制約を与えてしまう。プロセッサのクロックとデータを一緒に送信すれば動作周波数は向上できるが、プロセッサバスで頻発するプロセッサ間の切り替え時間は増加する。接続できるプロセッサ数は電気的に限界がある。また、バスの送信権を獲得するための遅れは別の性能限界を与える。さらに、プロセッサ間の転送の切り替えが頻繁で、プロセッサからの応答が一定しないプロセッサ間の高性能の接続には課題が残されている。

【０００５】プロセッサ間をスイッチで接続する場合、プロセッサとスイッチとの間で１方向に転送し、１対１の通信であることから動作周波数を向上できる。全体の通信容量は方式的には限界がなく、接続台数の制約はない。一方で、スイッチは外部に集中させる必要があるため接続はスター状になり、総配線量が多く、スイッチ周辺の配線が込み合う。また、全報知の通信は制御を複雑にして性能を劣化させる。スイッチを経由する遅延時間も性能を制約する。

【０００６】プロセッサ間の接続には、リングも用いられる。リング接続は、信号を一方向に１対１に転送すればよいので、バスより通信容量を大きくとれる。次のプロセッサまでの伝送線の長さが短く、クロック、ＬＳＩ、伝送線のスキューの補正を可能にし、信号の引出線が不要であり、次のプロセッサだけを駆動すればよいので動作周波数の制約を軽減できる可能性がある。また、バスのような接続の長さの制約は少なく接続台数の電気的制約もない。バスと同じく付加回路は必要なく、バスと同様に接続の配線量は少ない。さらに、すべてのプロセッサはリングの内容を見ることができるため、密結合のプロセッサ間接続の全報知の通信（スヌープ）にも適する。

【０００７】一方で、リング接続の場合、バスに対して２倍のピン数を必要とする。バスでは１クロックですべてのプロセッサに送信できるが、リングではプロセッサを経由する時間遅れがスキュー補正、クロック同期、送信権の制御等のため大きく、特に密結合のプロセッサ間接続には課題がある。

【０００８】プロセッサ間に応用したリングとして、例えば、ＩＥＥＥ Ｓｔａｎｄａｒｄ１５９６－１９９２，“Ｓｃａｌａｂｌｅ Ｃｏｈｅｒｅｎｔ Ｉｎｔｅｒｆａｃｅ（ＳＣＩ），”１９９２、あるいは、Ｄ．Ｃｅｃｃｈｉ，Ｍ．Ｄｉｎａ，Ｃ．Ｐｒｅｕｓｓ，“Ａ１ＧＢ／Ｓ ＳＣＩ Ｌｉｎｋ ｉｎ ０．８μＭＢｉＣＭＯＳ，”１９９５ ＩＥＥＥ Ｉｎｔｅｒｎａｔｉｏｎａｌ Ｓｏｌｉｄ－Ｓｔａｔｅ Ｃｉｒｃｕｉｔｓ Ｃｏｎｆｅｒｅｎｃｅ，Ｓａｎ Ｆｒａｎｃｉｓｃｏ，ＣＡ，Ｄｉｇｅｓｔ ｏｆ Ｔｅｃｈｎｉｃａｌ Ｐａｐｅｒｓ．Ｐａｐｅｒ ２０．２，Ｆｅｂｒａｒｙ，１９９５，ｐｐ．３２６－３２７．に記載のものがある。

【０００９】上記文献に記載されている代表的なプロセッサ間接続であるＳＣＩは、上述の課題を十分には解決していない。

【００１０】プロセッサ間接続手段としてリング接続を用いた場合のリングのインターフェース部の構成を図２に示す。

【００１１】図２において、バッファ８０１から受信して、フラグ及びアドレスを判定して受信ＦＩＦＯ８０７、或いは中継用ＦＩＦＯ８０６に送る処理をする。受信クロックで動作する範囲８０２とプロセッサ等の当該モジュールのクロックは違うものを使用するので速度差吸収用のバッファである受信ＦＩＦＯ８０７、及び中継用ＦＩＦＯ８０６が必要である。受信クロックと当該モジュールのクロックは当然位相も異なるのでバッファ８０１と送受処理８０５の間にクロックの乗り換えが必要である。送信時には転送ラッチ８０３を経由して送信する。なお、時々の同期パターンによってバッファ８０１に入る前に一括してスキュー補正される。送達確認のパケットは、自動的に生成されて送信される。リングは１つであり、すべての情報交換は同一のパケット形式によっている。スヌープなど全報知の機能を用いないでキャッシュコヒーレンスを行う。リングの障害検出、切り離し、復旧は、通常のパケットの転送パスを通じて行う。

【００１２】

【発明が解決しようとする課題】従来技術に述べた通り、プロセッサバスに適合するためにはレイテンシ、スループット、運用性を改善し、またプロセッサバス固有の機能を効率よく実現する必要がある。そして、リング接続の場合のプロセッサを経由するための時間遅れを改善し、比較的近距離の接続を前提にするプロセッサ・バスの性能限界を超えることを可能とする必要がある。

【００１３】本発明の目的は、プロセッサを経由するためのレイテンシを短縮することを可能とするリング接続を用いたデータ転送方法及び情報処理システムを提供することにある。

5

【００１４】本発明の他の目的は、リングが通常動作しなくなる切り替え時にもリングよりモジュールに安定したクロックを供給することを可能とするリング接続を用いた情報処理システムを提供することにある。

【００１５】本発明の他の目的は、情報処理システムのスループットを向上させるリング接続を用いたデータ転送方法及び情報処理システムを提供することにある。

【００１６】本発明の他の目的は、リングの構成変更に伴うシステム停止時間を短縮するリング接続を用いたデータ転送方法及び情報処理システムを提供することにある。

【００１７】

【課題を解決するための手段】

［プロセッサバスへの適用］上記目的を達成するため、本発明のリング接続を用いたデータ転送方法は、複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおいて、送信権を獲得するためのフラグ用の独立する信号線を設け、前記フラグをセットして送信権を要求し、受信したフラグがセットされていないことから前記送信権を先行して獲得したことを後から確認することで受信及び送信動作と前記リングの転送動作を並行して行う。

【００１８】別の観点からは、本発明のリング接続を用いたデータ転送方法は、複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおいて、同一クロックを前記リングに配信し、転送及び送受信に使用する。

【００１９】さらに別の観点からは、本発明のリング接続を用いたデータ転送方法は、複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおいて、位相調整回路によりクロックの位相を受信する信号に合わせ、クロックの乗り換えの遅延を不要にするためクロックを転送制御、受信制御、送信制御に同一位相で使用し、前記複数のモジュールのうちの少なくとも一つのモジュールは回送モードで前記リング１周の遅延を前記クロックの整数倍のサイクルに合わせる。

【００２０】さらに別の観点からは、本発明のリング接続を用いたデータ転送方法は、複数の信号線からなるリングでプロセッサ等のモジュールを複数接続した情報処理システムにおいて、複数のモジュールのうちの各モジュールはラッチせずに次のモジュールに転送し、ラッチ前の位相に合わせて送信する。

【００２１】さらに別の観点からは、本発明のリング接続を用いた情報処理システムは、複数の信号線からなるリングでプロセッサ等のモジュールを複数接続し、送信権を獲得するためのフラグ用の独立する信号線を設け、前記フラグをセットして送信権を要求し、受信したフラグがセットされていないことから前記送信権を先行して獲得したことを後から確認する手段を有する。

6

【００２２】さらに別の観点からは、本発明のリング接続を用いた情報処理システムは、複数の信号線からなるリングでプロセッサ等のモジュールを複数接続し、クロックの位相を受信する信号に合わせる位相調整回路と、転送制御、受信制御、送信制御に同一位相で使用するクロックを供給する手段と、前記リング１周の遅延を前記クロックの整数倍のサイクルに合わせる少なくとも一つの回送モードのモジュールを有する。

【００２３】さらに別の観点からは、本発明のリング接続を用いた情報処理システムは、複数の信号線からなるリングでプロセッサ等のモジュールを複数接続し、複数のモジュールのうちの各モジュールは、ラッチせずに複数のモジュールのうちの次のモジュールに転送する手段と、ラッチ前の位相に合わせて送信する手段とを有する。

【００２４】さらに別の観点からは、本発明のリング接続を用いた情報処理システムは、複数の信号線からなるリングでプロセッサ等のモジュールを複数接続し、複数のモジュールのうちの各モジュールは、ラッチせずに前記複数のモジュールのうちの次のモジュールに転送する手段と、ラッチ前の位相に合わせて送信する手段とを有する。

【００２５】さらに別の観点からは、本発明のリング接続を用いた情報処理システムは、複数の信号線からなるリングでプロセッサ等のモジュールを複数接続し、パケットを連続的に転送するブロック・パケットの送信権を獲得するための信号線と、送信権を獲得した２サイクル後からブロック転送の間通常のパケットの送信要求を抑止させことを通知する信号線と、ブロックパケットの送信権の獲得後の４サイクル後からブロックパケットの送信を許可する手段とを有する。

【００２６】

【発明の実施の形態】以下、本発明の実施例を図面を参照して詳細に説明する。

【００２７】最初に、本発明の基本的な概念を説明する。

【００２８】［リングのプロセッサバスへの適用］本発明においては、高速の遠距離の伝送手段として利用されているリングをその高速性を活かしてバスの性能限界を超えるためにプロセッサバスに適用する。プロセッサバスの最も重要な性能要素は遠距離の伝送では余り重要でないレイテンシの短縮である。一方、リングは転送の容量を共有するので本質的にスループットのネックとなる。また、遠距離の伝送では伝送系の障害回復が重要で、伝送系は独立している。それに対して、プロセッサバスでは接続モジュール数が少ない近距離の伝送であり、クロック、電源、冷却、筐体を共有する等から伝送系の障害は少ない。したがって、モジュールの障害検出、切り離し、構成変更等リングを含むシステムの一体運用が重要であり、レイテンシ、スループット、運用性

を改善してリングをプロセッサバスに適用する。

【００２９】［１サイクル１パケット転送］１サイクル１パケット転送によりレイテンシの３サイクル短縮を図る。現状のリングは遠距離の伝送を前提にしており、転送のビット幅は距離の制約から１ビットから１６ビット程度である。図３に示すように転送のビット幅が１６ビットで１パケットが６４ビットならば４サイクルになる。本発明のリングは近距離の伝送を前提にし、1サイクルで１パケットを１度に転送する方式をとる。図４に示すように１パケットを６４ビットにすると転送のビット幅は６４ビットで１サイクルになる。転送の方式に柔軟性を持たせるため、後述するようにブロックパケット、複数サイクルのパケットの転送も可能に拡張している。

【００３０】［送信権の先行獲得］送信権の先行獲得によりレイテンシの１サイクル短縮を図る。現状のリングは図５に示すようにフラグを読みその内容に従い送信権を獲得してフラグをセットして次のモジュールにフラグを送信する。本発明のリングは、送信したいならばフラグをセットして図６に示す送信、論理和、転送ラッチを経由して次のモジュールにフラグを送信する。送信権を先行して獲得したことを受信したフラグがセットされていないことで後で確認する。

【００３１】送信したモジュールはリング１周の時に合わせて、受信したフラグの転送を送信権を放棄との論理積により抑止してフラグをリセットする。各モジュールは図７に示すようにすべての情報を受け取りパケット内のヘッダを解読して自分宛のパケットだけを選択して受信し、並行してすべての情報を転送する。この方式により、読み、処理、書くという処理とパケットの転送を独立、並行させられる。パケットの送信は、あらかじめ送信権を獲得してその２サイクル後に行う。パケットを次のモジュールに単に転送すれば良いのでレイテンシを改善できる。

【００３２】［同一クロック配信］同一クロック配信によりレイテンシの３－４サイクルの短縮を図る。現状のリングは、各モジュールが遠距離にあることを前提にしているので独立するクロックを使用しており、図８に示すようにクロックの速度差の吸収のためのバッファがいる。本発明では、制御リングによりクロックを配信することにより速度差の吸収バッファを不要にする。

【００３３】［受信位相での処理］受信位相で処理することによりレイテンシの２分の１サイクル短縮を図る。現状のリングは、クロックの位相がモジュール毎に独立しているので、速度差吸収バッファに加え、図９に示すようにモジュール内で使用する時にクロックの乗り換えの遅延がいる。本発明では、図１０に示すようにクロック位相調整回路により受信するタイミングを合わせ、これを処理、送信に使用するのでクロックの乗り換えが不要になる。

【００３４】このようにすることにより、動作周波数はクロックのスキューに影響されなくなるので向上する。クロックの位相は、モジュールを経由する毎に遅延量に応じて進む。リング１周して再送信するときにはどのモジュールにおいてもクロックの整数倍になる必要がある。図１１に示すように、リングに最低一つのモジュールは、回送モードを設けてリングの１周のサイクル数をクロックの整数倍（ここでは６サイクル）にするため、クロックに同期して送信する。本発明では、伝送系の変動があまりないことを前提にしており、動作中にテストパケットを送信して位相の調整を動的に行うのではなく、エラー発生時のみ再調整する方法を採用する。

【００３５】［位相変化時のクロックの切り替え］位相変化時のクロックの切り替えによりクロックを連続供給する。モジュールの挿抜時、イニシャライズ時には、クロックの波形変動による影響を避け、連続的に動作を継続するため、ＰＬＬを経由してクロックを配信して、リング制御に関係する部分以外のモジュールの制御に用いモジュールの動作を保証する。図１２に本発明のクロック系統を示す。

【００３６】［ラッチせず中継する方式］ラッチせず中継することによりレイテンシを２分の１サイクル短縮する。現状のリングでは図１３に示すようにラッチを経由して受信パケットを転送する。本発明のリングでも、波形整形のため転送ラッチを経由して受信パケットを転送する。本発明では、図１４に示すように、波形整形の必要がない場合は転送ラッチを経由せずに中継できる。ラッチは受信の立ち上がりの２分の１のタイミングで行うのでこれにより２分の１サイクルのレイテンシの短縮ができる。

【００３７】送信パケットの位相は、転送ラッチで位相の調整ができないので、受信の立ち上がりに位相を合わせるためのラッチを経由して送信する。本発明の各モジュールは、ラッチせず中継する中継モード、波形整形のため数モジュールに１つ必要になる転送ラッチを経由して中継する転送モード、リングに最低１つ必要なクロックに同期させる回送ラッチを経由する回送モードのいずれかのモードで動作する。図１５にパケットＡ、Ｂ、Ｃを転送する例を示す。

【００３８】［受信線毎のスキュー補正］受信線毎のスキュー補正により動作周波数を向上させる。現状のリングは、スキューを一括して補正してラッチしているので動作周波数はスキュー量が直接的に影響する。本発明では、簡単な回路により図１６に示すように信号線毎にスキュー補正するので動作周波数は補正精度で決まる所まで向上できる。

【００３９】［信号線を用いた送達確認］信号線を用いた送達確認によりスループットを向上させる。現状のリングは、送達確認を応答パケットによっている。本発明のリングは、応答パケットを送信する代わりに送達確認

のために送達確認とその報告を信号線で行うのでスループットを改善できる。送達確認をパケット毎に行う方式を取れば、スループットは２分の１になる。送信モジュールは、送信権獲得時に到達確認の報告と通知の２つの信号線の送信権を獲得する。受信モジュールは、パケットの受信の２サイクル後に送達確認を信号線で報告する。全報知についても送達確認を行うために受信できた時には何も応答しない応答方式を取るので、信号線１本で送達確認ができる。

【００４０】［短いデータ幅のリングの併設］短いデータ幅のリングを併設することによりスループットを向上させる。現状のリングは、遠距離の転送を目的としてきたため、用途別に短いデータ幅のリングを併設するという概念はない。短いパケットもリングを共用していた。プロセッサバスに適用するとデータ転送以外の短いパケットが多量に必要になる場合がある。データ転送と共用するとスループット、転送効率が下がる。本発明のリングは、信号１本で送信権の制御ができるのでブロックパケットの転送を行うブロックパケットリング、１サイクルで２つのパケットを並列転送するための選択リング、キャッシュのディレクトリの読み出すためのディレクトリ・リング等短いデータ幅のパケットを持つリングを容易に併設できる。データ転送のためのリングを基本リングと呼ぶ。図１７にリングの種類を示す。基本リングと併設リングはクロックを共通にすることにより、情報の交換、相互の動作の同期を行う。制御リングも共用するのでリング併設による複雑性を除去する。

【００４１】［ブロックパケットリング］ブロックパケットリングによりスループットを向上させる。ブロックパケットリングは、パケットを連続的に転送して各モジュールの制御を簡単にし、またスループットを向上させるためにある。現状のリングでは、パケットのヘッダに連続転送のフラグに相当するものがある。この方式を取ると受信パケットを読み、処理、書くことで送信権を獲得し、通知することになり本方式で述べたレイテンシの短縮ができない。本リングの特徴を活かして、基本リングを拡張してブロックパケットを転送できる。

【００４２】基本リングは、１サイクルに１パケットを転送している。まず、パケットの連続転送するモジュールを１つ選択する必要がある。選択のための送信権を獲得する機能を設ける。パケットの連続転送中は１サイクルに１パケットの転送を抑止しなければならない。これには、あらかじめパケットの連続送信を各モジュールに通知すればよい。各モジュールは、予約を通知された２サイクル後からパケットの送信要求を抑止する。かくして、送信モジュールは連続送信を通知した２サイクル後からパケットの送信要求を行うことができる。このため、図１８に示すようにパケットの連続転送の送信権獲得のための信号である予約線、及び予約通知線の２本の信号からなるブロックパケットリングを併設すればパケットの連続転送を実現できる。

【００４３】［選択リング］選択リングによりスループットを向上させる。基本リングは、送信権をリング１周の間獲得する。送信モジュールから受信モジュールまでパケットを転送する場合、受信モジュールから送信モジュールまで同時にパケットを転送できる。図１９に示すようにモジュール１（１０４）からモジュール２（１０５）の転送１０６とモジュール３（１０１）とモジュール０（１０２）との間の転送１０３は同時にできる。これが実現できれば、スループットを最大２倍に向上することができる。リングにおいてこれを実現した例は見あたらない。

【００４４】従来の一般的な技術として、各モジュールから並列的に要求を集め選択する集中的アービトレーションの方式が知られる。この部分は、スイッチと同じトポロジーになる。集中的アービトレーション方式は共通部分が必要でリングの拡張性を減ずる。集中的アービトレーションをそのままリングで実現すると各モジュールからのパケット数は送信要求と同じになり、リングのスループットの大半を費やす。しかも送信要求を各モジュールが同時に行うには接続モジュール数倍の信号線が必要になる。

【００４５】本リングの特性を活かして集中的アービトレーションをリングで実現することができる。ブロックパケットの転送には数サイクル掛かるので、送信要求をモジュール全体に対して１サイクルに１つ行う能力があればバランスが取れる。これにより送信要求を送るための送信権を獲得する機構を全体で１つ設ければ良いので、信号線はモジュール１つ分で良い。送信権を獲得したモジュールは、送信要求をアドレス０のモジュールに送る。アドレス０のモジュールは、各モジュールからの送信要求を同時に最大２つ選択して送信権を通知する。送信要求は、要求アドレス、宛先アドレスからなり、要求アドレスで送信権を通知する。信号を追加すれば、ブロックパケットの要求、ビジーの報告も行うことができる。基本リングの送達確認の信号線２本に加え並列転送のため、もう１組を追加する。

【００４６】［並列送信要求選択方式］従来の集中的アービトレーションでは１サイクルに１つの送信要求を選択している。本リングでは、１サイクルに２つの送信要求を選択すること、要求アドレス、宛先アドレスから転送可能性を判定することなど、従来技術にない処理能力が要求される。本発明のリングは、図２０に示すように、送信要求をモジュール毎にバッファし１つ目の送信権を転送可能なブロックを並列的に判定してその中から最初のものを選択し、２つ目の送信権を最初の送信権と並行転送可能なブロックを並列的に判定してその中から最初のものを選択する高速論理方式である並列送信要求選択方式を用いている。

【００４７】［ディレクトリ・リング］各モジュール毎

に、それに属するキャッシュにあるラインの状態のみを格納するディレクトリを持ち、１カ所に集中してディレクトリを持たない分散方式がある。この分散方式では、各モジュールに属するディレクトリよりキャッシュの状態を集約し、自分のモジュールに属するディレクトリの状態を決め、メモリモジュールに通知する必要がある。この様な場合の従来のキャッシュの状態制御は、状態の報告を各モジュールから受け、集約してメモリモジュールに配信する集中方式を用いている。この部分はトポロジー的にはスイッチと同じになる。集中的キャッシュ状態制御には共通部分が必要でリングの拡張性を減ずる。キャッシュの状態を調べる時間は一定でなく、キャッシュの状態の報告は、全モジュールから受け取ったことを確認する必要がある。従来の方式をそのままリングに応用すると全モジュールからキャッシュの状態をパケットで報告することになる。これでは、送信要求に対してそのモジュール数倍のパケットが必要になりリングのスループットの大半を消費する。

【００４８】ディレクトリ・リングを用いることにより、本発明のリングの特性を活かしてキャッシュの状態制御をディレクトリ・リングで実現できる。報告を求めるモジュールから全モジュールについて報告を求めることができればパケット数は減らせる可能性がある。報告を求める間隔を報告の揃う時期に設定すれば高々１－２パケットで全モジュールからのキャッシュの状態の報告を受け取れる。パケットに各モジュール毎に状態を報告する領域を設けると、モジュール数倍の信号線を要する。キャッシュの状態の種類別に情報線を定めて集約できれば信号線が少なくて良いが、どのモジュールが報告したかの情報は失われる。

【００４９】必要な機能は、全モジュールが報告したことを確認することである。この確認のためにキャッシュの状態を報告していないことを示す信号線を１本設ける。この信号線でキャッシュの状態報告できないことを知らせる。この信号線が有意でなければ全モジュールが報告したことになる。かくして、図２１に示すようにキャッシュの状態報告を要求するための送信権の獲得、状態報告の要求、キャッシュの状態報告、未確定の報告、集約したキャッシュの状態の通知、確定の通知に必要な信号線を設けることによりリングで実現できる。未確定報告と同時に、その種類別に未確定原因を報告すれば、要求モジュールは効率よい再要求のタイミングを選択できる。

【００５０】［階層接続］階層接続によりスループットを向上させる。ローカルメモリのキャッシュコヒーレンスを行うリングとして従来技術として前述したＳＣＩがある。ＳＣＩにおいては、キャッシュコヒーレンスを全報知しないことを前提にパケットの問い合わせ、応答を階層的に行うためにパケット数が多くなり、確定までの時間が掛かる。また、ローカルメモリに対する完全なディレクトリを必要とするために容量も大きい。

【００５１】本発明のリングは、ローカル・アクセスには性能の劣化がなく、リモート・アクセスには単純なディレクトリを設けて必要なモジュールのみにアクセスする手段を提供する。ディレクトリにエントリがないときにはリングの全報知の高性能を活かして全報知する方式を採用してレイテンシの改善と簡素化を行う。リングに属するローカルメモリのラインがリングの外で変更されているかが分かればリングの外にアクセスするかを決められる。リングに属するローカルメモリのラインがリングの外にコピーがあればコヒーレンスの動作をするかを決められる。リングに属するローカルメモリの全ラインに対して他のリングに変更有り、コピー有りの２ビットを持つディレクトリ（ローカル・ディレクトリ）を設ければ容量を余り必要とせずにローカルアクセスについて他のリングをアクセスすることはなく性能劣化はない。変更或いはコピーのあるリングを特定する情報をディレクトリに追加すると必要な他のリングにのみアクセスすれば良いので性能は上がるがディレクトリの容量は大きくなる。ローカルメモリの全ラインに対してでなく最新のアクセスのラインを登録するディレクトリ（リモート・ディレクトリ）を設け、各リング或いは各リングのグループでの変更、コピーの状態のベクトルを持てば容量を小さくできる。変更のアクセスは排他制御のために頻発するのでディレクトリに優先して維持する。リモート・ディレクトリにエントリがあれば該当するリングだけにアクセスする。リモート・ディレクトリにエントリがなければ、全リングにアクセスする。図２２に示す様にいずれの場合も、アクセスのパケットと兼用でき、１回のパケットの転送でコヒーレンスを取れるので時間は短い。キャッシュコヒーレンスの時間を短縮することでレイテンシ、全報知パケットを削減することでスループットをＳＭＰ並に改善することができる。ローカルメモリのアクセスの抑止は、ローカルメモリにもディレクトリを持つか、ローカルメモリを直接制御すれば、階層接続機構を通じて知らせる必要はない。その方法がないときには、ディレクトリが分散しているときの手段であるディレクトリリングを設ける必要がある。

【００５２】［ディレクトリのローカルメモリへの格納］ディレクトリのローカルメモリへの格納により処理効率を向上させる。ローカル・ディレクトリの容量が大きくなるときには、ローカルメモリにローカル・ディレクトリを置きその写しを階層接続機構に持つ。階層接続機構が受信不能である時にディレクトリのアクセスのためにローカルメモリをアクセスするとその応答を受け取れない。このデッドロック状態を解決するために、従来技術である論理チャネル、或いはサブチャネルの概念を一部導入する。リングのアドレスを物理チャネル或いはチャネルと見立てて、リングの本来のアクセスとディレクトリのアクセスをそれぞれ論理チャネル０、１と解釈

する。この場合に必要になるのは、どの論理チャネルがパケットを受け取れるかであるので論理チャネルのステータスに相当するものとして受信不能の報告時に論理チャネル０、１を指定する応答信号を１本追加する。各モジュールは、フロー制御、ビジー状態をモジュールアドレスと論理チャネル０、１で管理する。論理チャネルの拡張、ステータスの追加には応答信号の本数を増やす。

【００５３】［制御リング］現状のリングでは、モジュールの障害検出、特定、切り離しはリングの通常動作を利用して行うので、そのための論理量が多い。障害モジュールの特定、切り離しのためには図２３に示すように逆方向に回る２つのリングが必要になる。現状のリングは伝送系の障害が支配的である遠距離の伝送を前提とするため、この２つのリングは折り返し機能を実現する冗長系として効果がある。しかし、冗長系にするとレイテンシは、２倍になりリングの構成要素は２倍になる。プロセッサバスではモジュールの障害に起因するリングの障害が支配的で、伝送系に起因する障害はむしろ例外的である。プロセッサバスでは、リングに接続するモジュールの障害に起因するリングの障害検出、特定、切り離し等のシステム運用が重要である。

【００５４】本発明のリングは、障害検出、特定、切り離しの論理量を小さくして、構成変更、障害検出、特定、切り離しを独立して行うリング制御ユニットを設け、低速の制御データ線と制御クロックの２本の信号線とクロックからなる制御リングで各モジュールに接続する。リング制御ユニットからモジュールには転送によらずに配信する方式として、制御リングのモジュールの構成変更、障害に伴う各モジュールへの影響を軽減する。

【００５５】図２４に示すように、データ転送のための１つのリングと３本の信号線からなる制御リングにより、冗長リングを設けるよりも低コストで運用性の高いリングを構成できる。制御リングは直接的な制御をするので、各モジュールでの運用機能は不要になる。モジュールは障害を直接に報告できるので障害の検出の論理量が少ない。切り離しは障害モジュールにリングの中継を指示するだけなので切り離しのための論理量は少ない。制御リングにより、クロックの位相調整、受信線の遅延量をイニシャライズするので、各モジュールのイニシャライズ機能は簡素化できる。切り離しできない障害は、ケーブル、コネクタ、中継ゲート、ドライバ、レシーバに限定される。モジュールの挿抜に伴うリングの切断、復旧をモジュールから制御リングを経由して行うので時間を短縮できる。モジュールの内部には、イニシャライズ中もクロックを供給してイニシャライズ終了後位相調整されたクロックに連続的に切り替えるので各モジュールは連続動作できる。

【００５６】［構成変更の簡便性］従来のリングにおいて、中継部分を含む接続装置の追加、除去はケーブルの接続変更を要する。プロセッサリングにおいては、装置の追加、除去に相当するモジュールの挿抜を行うには、短絡スイッチとの入れ替えになる。その際、モジュールの引き抜きと短絡スイッチの挿入と２つの動作が必要である。本発明では、冗長リングを用いずに構成変更の簡便性を上げる。モジュールの挿抜に伴い、その力を使用して連動して挿抜する短絡スイッチを考案してプロセッサバスへの適用に伴う構成変更の時間の短縮と簡便性を実現した。

【００５７】モジュールの挿抜を動作中に前触れなく行うとパケットの送信と応答など一連の動作が中止されシステムは異常に停止する。一般には、モジュールの挿抜を行う前にオペレータによりシステムを停止し、挿抜後にシステムを立ち上げるのでシステムの停止時間が長い。そこで、モジュールの挿入を事前に検知するセンサを設け、次のモジュールにセンサを接続して制御リング経由で通知する機能を設ける。また、モジュールの引き抜きを事前に検知するセンサを設置し制御リング経由で通知する機能を設ける事前に挿抜するモジュールだけを切り離して、その後は必要最小限の時間だけ自動的にシステムを停止すれば良いので、業務処理の停止時間を短縮することができる。

【００５８】［論理的順序づけ、双方向リング］本発明のリングは、メモリを用いるセマフォアのための論理的順序づけを保証できる。また、双方向リングを使えば、スループットを約２倍に向上できる。

【００５９】［レイテンシ］図２５に本発明のリングによるレイテンシの改善内容を示す。

【００６０】［動作周波数］図２６に本発明のリングによる動作周波数の改善を示す。

【００６１】［増設性］本発明のリングは、同一周波数、リング配線での送信権リングの追加、階層接続により性能を向上することができる。また、同一周波数、リング配線での双方向リング化による性能向上とクラスタ化ｈが可能である。クラスタ化にいたる増設性のパスと同一のＬＳＩ技術を用いた性能向上の一例を図２７に示す。

【００６２】［単一リングの性能］動作周波数の限界は、送信ラッチの出力毎にスキュー補正をしない場合は、スキュー補正回路の精度によるスキューと送信ラッチのスキューの重畳で決まる。送信ラッチの出力毎にスキュー補正をすればスキュー補正回路の精度だけで決まる。ドライバの周波数特性は一般にこれより大きいので、概ねドライバの動作周波数から考えて４００Ｍｈｚ程度と考えられる。

【００６３】レイテンシを８モジュール、配線長の最大１０ｃｍで考える。中継モード のモジュールのレイテンシは、ゲートディレイ　２段（０．３ｎｓ）、ドライバの遅れ（０．８ｎｓ）、配線（１０ｃｍ）を含む実装遅れ（１ｎｓ）とスキューの平均遅れ（０．４ｎｓ）とで、合計２．５ｎｓである。転送モードのモジュールの

レイテンシは、中継モードのモジュールのレイテンシに加え、ラッチの受信ウィンドの中央（平均０．４サイクル、４００Ｍｈｚの時、１ｎｓ）とラッチの遅れ（０．５ｎｓ）で、合計４ｎｓである。

【００６４】中継モードのモジュールが６、転送モードと回送モードが１であると、リング１周の平均レイテンシは　２３．５ｎｓになる。リングの一周のサイクル数は、４００Ｍｈｚの場合１０サイクルになる。２５０Ｍｈｚの場合、転送モードと回送モードのレイテンシがそれぞれ０．８ｎｓ増えるのでリング１周の平均レイテンシは　２５．１ｎｓ、リング１周のサイクル数は７サイクルになる。モジュール間の平均レイテンシは、ほぼリングの半周になるので４００Ｍｈｚの場合、約１２．５ｎｓ、２５０Ｍｈｚの場合１４ｎｓになる。

【００６５】［階層接続の性能］階層接続では、（１）送信副リングでの送信モジュールから接続機構、（２）送信接続機構の処理、（３）主リングでの送信接続機構から受信接続機構、（４）受信接続機構の処理、（５）受信副リングでの接続機構から受信モジュールまでがレイテンシになる。

【００６６】（１）、（３）、（５）は、単一リングのレイテンシと同じであるので単一リングのレイテンシの３倍になり、２５０Ｍｈｚの場合４２ｎｓになる。

（２）の送信接続機構の処理は、受信と送信処理である。ディィレクトリのメモリへのアクセスはない。受信と送信にそれぞれ６サイクル、２５０Ｍｈｚの場合２４ｎｓになる。全報知のパケットは、ディレクトリをアクセスするのでこれを加える。（４）の受信接続機構の処理は、ディレクトリのアクセスと副リングへのパケットの転送を並行してできるので受信と送信に６サイクル、２５０Ｍｈｚの場合２４ｎｓとなる。これらの結果、全体では９０ｎｓになる。

【００６７】次に、本発明の実施例について説明する。

【００６８】［基本リング］最初に、基本リングについて説明する。

【００６９】［接続形態］対象とするプロセッサ間リングインターフェースは複数のプロセッサモジュール（ＰＭ）、入出力モジュール（ＩＯ）、メモリモジュール（ＭＭ）、リング制御ユニット（ＲＣ）等を接続する。ＰＭ、ＩＯ、ＭＭ合わせて８モジュールをリングに接続する。ＰＭ、ＩＯ、ＭＭの各モジュールの信号は、基本リングとして選択線（１本）、アドレス線（３本）、データ線（７２本）、制御リングとしてクロック（１本）、制御クロック（１本）、制御データ線（１本）である。各モジュールは、基本リングの選択線、アドレス線、データ線を前のモジュールから受信し、次のモジュールに転送する。リング制御ユニットは、基本リングを動作させる高速のクロック、基本リングを制御する低速の制御クロック、制御クロックに同期した制御データを発生し配信する。

【００７０】モジュールには、回送モードと転送モードの動作モードがある。リングに１つのモジュールは選択線、アドレス線、データ線をクロックに同期して転送する回送モードで動作する。基本リングの１周は、クロックの整数倍の遅延で２サイクル以上に設定する。図２８に基本リングの各モジュール間の接続を示す。

【００７１】各モジュールは、制御リングの制御クロックで制御データを処理し、自分が宛先であるときに受信する。内容に応じて、モジュールの状態についての応答を制御データに乗せて送信する。リング制御ユニットは、各モジュールにポーリングして各モジュールの情報を受信する。クロックは、動作周波数の２分の１の周波数を配信してモジュールで動作周波数に倍周する。クロックの動作周波数の制約は軽減される。以下の説明では簡単のため動作周波数のクロックを供給することで説明する。図２９にリング制御ユニットに関係するリングインターフェースの信号を示す。

【００７２】［送信権の先行獲得］パケットは要求元、要求番号及び種別等のヘッダと本体とからなるが、その形式及び内容についてはリング固有の課題はないのでここには述べない。送信権については、選択線、アドレス線を受信しその内容に従い獲得する方式がよく知られる。図３０に従来の方式を適用する例を示す。選択線７２０は入力ラッチ７２１に格納され、送信要求７２２について送信権処理７２４を行い、結果を出力ラッチ７２３に格納する。レイテンシは最低１サイクルは生じる。このレイテンシを短縮する方式をここに述べる。

【００７３】図３１に、選択線をゲート２段の遅れで転送できる方式を示す。送信要求７２７はこのまま選択線７２０と論理和７２８をとり転送する。選択線７２０の内容と送信要求７２７をしていることから送信権獲得確認７２６をする。選択線は、自分の送信権を放棄（７２５）する時には選択線を論理積７２９で抑止する。

【００７４】図３２に受信線についての従来の方式を適用する例を示す。受信線７３３は受信ラッチ７３０に格納され、送受信処理７３１を行い送信ラッチ７３２に格納する。レイテンシは、最低１サイクルは生じる。

【００７５】図３３に、受信処理７３４と送信処理７３５を並行して行う方式を示す。受信線７３３は受信処理（７３４）され、並行して送信処理７３５を行い、送信の場合は論理積７３７により受信線を抑止する。

【００７６】パケットの宛先はアドレス線を用い、パケットはデータ線を用いる。なお以下の説明で明確になるが、アドレス線は説明のため分離しているが、実際にはパケットのヘッダに含まれるので実際の信号線はデータ線に含まれる。データ線（７２本）の内８本はパケットの内容に応じてエラー訂正符号或いはパリティに用いられる。実際のデータは、６４本（６４ビット）である。宛先とパケットはリング上のクロックに同期した１サイクルで転送する。選択線は２サイクル後のデータ線、ア

ドレス線にパケットと宛先があることを示す。

【００７７】図３４は、時間２（２０１）で選択線を１（２０２）にしてその２サイクル後の時間４（２０３）でアドレス線に宛先３（２０４）を、データ線にパケットＡ（２０５）を、時間４（２０３）で選択線を１（２０６）にしてアドレス線に宛先５（２０７）を、データ線にパケットＢ（２０８）を送信することを示す。

【００７８】［受信］信号の変化の後縁でラッチからラッチに転送する受信方式を図３５を用いて説明する。受信１０は当該モジュールが送信をせず受信状態にあることを示す。前のモジュールからのデータは、データ線１１を経由して当該モジュールに転送される。受信１０が１であるので論理積１２が成立し、データ線１１は論理和回路１３を経由して次のモジュールに転送する。データ線の内容は受信ラッチ１４に常にラッチする。アドレス線についても同様に受信ラッチ１５にラッチし当該モジュールの自分アドレス１６と一致回路１７で比較する。

【００７９】選択線の内容が１であれば、２サイクル後のアドレス線とデータ線に宛先とパケットがあることを示す。選択線はラッチ１８、ラッチ１９、ラッチ２０に１サイクルづつ遅れてラッチされる。ラッチ２０の選択線の内容は受信ラッチ１５より２サイクル進んでいるので選択線の内容１でアドレスの一致論理１７が１ならば論理積２１が成立して、論理積２２を経由して受信パケット２３として受信する。図３６に受信動作をタイムチャートに示す。

【００８０】［送信権獲得と送信］送信したいパケットがあれば、選択線を１にする。受信した選択線が０であれば送信権を獲得し２サイクル後に送信モードとなる。受信した選択線が１であれば、送信権が獲得できなかったことを知る。そのまま受信を続け、かつ送信権の獲得待ちになり選択線が０になるのを待つ。選択線が０になれば、２サイクル後に宛先とパケットをアドレス線、データ線に乗せる。送信したパケットがリングを１周すれば選択線を０にし送信権を放棄する。そのパケットがリングを一周する所定のサイクルより２サイクル前迄に送信要求があれば引き続いて選択線を１にして送信権を獲得できる。

【００８１】図３７に送信動作の概念図を示す。送信要求６００があれば１（６０１）を出力する。受信した選択線６０２が１（６０３）ならば引き続き送信要求６００を１（６０１）にする。受信した選択線６０２が０（６０４）ならば送信要求を０（６０５）にして送信権を獲得したことを知る。選択線の出力（６０６）は、受信した選択線６０２と送信要求６００の論理和を出力する。

【００８２】図３８を用いて送信動作を述べる。当該モジュールが送信したいことは送信３０を１にすることで示されている。以前に送信３０を要求してそれが処理されていれば、要求（ラッチ）３１は０である。論理積３２が成立して受付（ラッチ）３３が１となる。受付３３は、送信３０が受理されたことを送信受付（３４）信号でモジュールの送信要求論理に知らせる。ついで新たな送信要求を調べ送信３０信号に伝える。毎サイクル送信要求をするには、受付３３の結果を用いて１サイクルで更新する必要がある。更新回路は、リングインターフェースの問題ではないのでここでは述べない。論理積３２が成立することにより同じく要求３１（ラッチ）が１になる。要求３１（ラッチ）の内容は直ちに論理和３５を通じて次のモジュールに伝える。当該モジュールが送信権を獲得後リング１周したサイクルであれば、送信権放棄３６は１になり論理積３７が成立せず選択線の転送を抑止する。

【００８３】他のモジュールが送信権を獲得していないサイクルならば選択線の内容は０であるので単に要求３１の内容を伝える。他のモジュールが送信権を獲得したサイクルならば選択線は１でありかつ論理積３７が成立してその送信権と当該モジュールの送信要求は論理和３５に重畳される。受信した選択線が０であれば、他のモジュールが送信権を獲得していないことを示し論理積３８が成立し獲得３９（ラッチ）が１になり送信権を獲得したことを知る。送信権放棄３６が１であるときは送信権の処理は選択線が１であっても選択線が０と同じ動作をする。送信権放棄３６が１であれば、当該モジュールの送信パケットがリングを１周したことを合い召し論理積４２が成立し獲得３９（ラッチ）が１になり送信権を獲得したことを知る。選択線が０になってから２サイクル後、獲得３９（ラッチ）が１になってから次のサイクルに送信モードとなりデータ線、アドレス線にパケットと宛先アドレスを転送する。受信した選択線が１であれば、論理積３８が成立せず獲得３９（ラッチ）は０であり送信権を獲得できなかったことを知る。送信権獲得待ちで選択線が０になると論理積３８が成立して獲得３９（ラッチ）が１になる。選択線が１で、要求３１（ラッチ）が１で送信権放棄３６が０のときには論理積４０が成立して要求を出し続ける。送信権の獲得持ちになり選択線が０になるのを待つ。選択線が０、要求３１（ラッチ）が１、且つ送信３０が１ならば論理積４１が成立して引き続いて送信権を獲得することになる。送信権放棄３６が１、要求３１（ラッチ）が１、且つ送信３０が１ならば論理積４３が成立して引き続いて送信権を獲得することになる。論理積４１、論理積４３が成立すると受付３３（ラッチ）が１になり送信３０を受付けたことを知らせる。獲得３９が１になると論理積４７が成立して送信ラッチ４８に宛先アドレスをラッチし論理和５０を経由して次のモジュールに送信する。獲得３９が１になると論理積５１が成立して送信ラッチ５２に送信パケットをラッチし論理和５３を経由して次のモジュールに送信する。ラッチ４４が１になるとアドレス線の内容を次

のモジュールに転送しないように論理積５４で抑止する。ラッチ４４が１になるとデータ線の内容を次のモジュールに転送しないように論理積５５で抑止する。送信論理は受信論理と同じく２段のゲート論理で構成される。

【００８４】図３９に送信動作の一例を示す。送信３０が１になり、要求３１が０であることから要求３１、受付３３が１になる。受付３３が１になることにより、送信３０が更新され１サイクルの中で０になる。選択線（入力）が引き続き１であるので要求３１は１である。受付３３は送信３０が０であるので次のサイクルは０になる。選択線（入力）が０になる次のサイクルで獲得３９は１になる。獲得３９が１になる次のサイクルでラッチ４４が１になる。ラッチ４４により送信ラッチ４８が更新しアドレス線に出力する。ラッチ４４により送信ラッチ５２が更新しデータ線に出力する。リング１周のサイクルである６サイクル後に送信権放棄３６が０となる。当該モジュールで送信要求が引き続きないので次のサイクルで選択線（出力）は０となる。

【００８５】［スキュー補正］従来の方式においては、クロックを受信線に対してスキューの設計値にセットアップ時間を加えただけ遅延させ、パルス幅はクロックのスキューがスキューの設計値だけ遅れても受信できる必要がある。動作周波数は、スキューの設計値の２倍にセットアップ時間を加えて決められる。図４０にこれを示す。

【００８６】図４１に受信線に合わせてクロックの位相を設定する方式を示す。動作周波数はスキューの設計値にセット時間を加えたものになる。レイテンシは、スキューの実績値とセット時間の和に短縮される。図４２に受信線毎にスキュー補正を行う方式を示す。動作周波数は、スキューの補正精度で決まる。

【００８７】［受信位相での処理］従来の方式ではリングのクロックと各モジュールのクロックの位相は独立である。平均すれば２分の１のサイクルの遅れを生じる。各モジュールは、受信線の位相に合わせて動作すればレイテンシを削減できる。リングに１つのモジュールはリング１周するサイクルをサイクルの整数倍にする必要がある。図４３に従来の方式を示す。リングのクロック７４０によって受信線７３３を受信ラッチ７３０に格納してモジュールのクロック７４２によりサイクル調整を行いモジュールのサイクルに同期する。送受信処理７３１を行い送信ラッチ７３２に格納して転送する。

【００８８】リングのクロックに同期して動作する方式を図４４に示す。受信線７３３は、クロック７４２を位相調整（７４０）して受信ラッチ７３０に格納し送受信処理７３１を行う。受信線７３３と送信内容を論理積７３７、論理和７３６で選択しリングクロック７４０で転送ラッチ７４３に格納する。各モジュールは、転送ラッチ７４３の出力を次のモジュールに転送する。これを転送モードと呼ぶ。リングに１つのモジュールは、転送ラッチ７４３をサイクル調整７４１を経てクロック７４２で回送ラッチ７４４に格納して次のモジュールに転送する。この同期方法を回送モードと呼ぶ。

【００８９】［転送モードの動作］選択線、アドレス線、データ線等の受信線のラッチの位相を受信線の遅延とスキューに合わせて転送ラッチのクロックの位相を定めて動作周波数、レイテンシを改善する。動作周波数には、クロックのスキューの設計値と受信線のスキューの設計値の和が直接的に影響を与える。クロックのスキュー補正を行えば動作周波数に与える影響はクロックスキューの設計値を加えずに、受信線のスキューの設計値だけでよいので２分の１にすることができる。クロックのスキュー自動補正を行えば、リングのレイテンシに与えるスキューの影響は、クロックのスキューの設計値を加えずに、受信線のスキューの設計値でなく受信線のスキューの実現値だけでよい。一般にスキューの設計値と実現値との乖離は大きいので動作周波数の改善に寄与する。

【００９０】図４５に転送モードの送信動作の概念図を示す。最も早く受信した受信線（６１０）と最も遅く受信した受信線（６１１）の中間の２分の１サイクル後に遅延クロック（６１２）を設定する。受信したデータは遅延クロック（６１２）で転送ラッチに格納して次のモジュールに転送する。受信した選択線（６１３）の位相は一般に両者の中間にある。受信した選択線６１３はラッチ（６１４）する。

【００９１】送信要求６１５があれば１（６１６）にする。ラッチ（６１４）した選択線６１３が１（６１７）であるので引き続き１（６１６）にする。ラッチ（６１４）した選択線６１３が０（６１８）になると送信要求６１５を０（６１９）にする。選択線６２０にはラッチ６１４と送信要求６１５の論理和を出力する。

【００９２】［位相調整回路］遅延クロックは、図１２に示すように倍周されたクロックから位相調整回路を経由して作る。図４６はゲートの遅延を利用した位相調整回路であって、半導体の特性ばらつきを考え、最小の遅れでも２分の１サイクルの遅延を保証できるように遅延量を取る。遅延量は遅延設定レジスタで指定する。遅延量は、選択回路１段、選択回路２段、選択回路２段にゲート２段からゲート１６段を加える８段階で合計１０段階、負クロックの入力をいれると２０段階になる。但し、半導体の特性ばらつきにより最大の遅れでは、クロックと負クロックの遅延が重複することがある。例えばゲート８段階で遅延が２分の１サイクルに達すると６段階になる。

【００９３】［調整クロック］調整クロックは、４サイクルの間位相を連続して変化する信号である。前のモジュールより設定用のパターンを送信して遅延設定を行う。前のモジュールは、信号の変化点を容易に特定でき

るように選択線、アドレス線、データ線に１０００の４サイクルの繰り返しパターンを送出する。設定用パターンの４サイクルに一つある信号の変化点を特定するために４サイクルの間連続的に位相を変化する調整クロックを設ける。調整クロックは、４サイクルの周期で１サイクル単位に位相を変化させる機構、位相調整回路、遅延がそのサイクルにあることの範囲検出回路を接続して発生する。１サイクル単位に位相を変化させる機構は、リングカウンタで４サイクルの周期の１０００の繰り返しパターンを発生し、サイクルを選択する機能を有する。上記信号とクロック及び逆極性のクロックとの論理積を取りどちらかを選択することで２分の１サイクル単位に位相を変化させる。上記の信号を位相調整回路の調整クロックに入力する。範囲検出回路は、位相調整回路の出力である遅延クロックと遅延前の調整クロックとの論理積を取りセットリセットラッチに入力してその出力信号を判定する。出力信号が１ならばその遅延の範囲が２分の１サイクル以下であると判定する。遅延がその２分の１サイクルの範囲を超えるか、位相調整回路の遅延量が最大になれば２分の１サイクル先の位相に移る。かくして、４サイクルの周期の可変遅延を有する受信位相決定用の調整クロックを作成できる。

【００９４】［受信位相の決定］受信位相の決定の手順を以下に示す。

【００９５】すべての受信線について設定用パターンの０を受信する調整クロックの位相を定める。少なくとも１つの受信線について１を受信する遅延クロックの位相を定め、前縁とする。すべての受信線について１を受信する遅延クロックの位相を定め後縁とする。前縁と後縁の中間の２分の１サイクル後に受信位相を定める。前縁と後縁の中間は次の手順による。前縁の検出時に計数を開始して後縁で終了するスキューカウンタを設ける。次に前縁に調整クロックを設定する。スキューカウンタの２分の１迄調整クロックの位相を進めて前縁と後縁の中間に位相調整回路の遅延量を定める。前縁と後縁の中間の２分の１サイクル後のクロックは、調整クロックを作成する元のクロックとは逆極性のクロックを位相調整回路に入力することで得られる。

【００９６】［中継モード］中継モードでは、転送ラッチの入力を受信データとして転送する。転送ラッチの入力の切り替わりは転送ラッチのタイミングの２分の１サイクル前である。転送モードの送信では、送信内容は転送ラッチのタイミングで切り替わっていて、１サイクル後に転送ラッチにラッチしてから転送している。中継モードの送信では、１サイクル前の送信内容を転送ラッチの入力の切り替わりに合わせるため２分の１サイクル遅延させて中継する必要がある。このため、送信内容を転送ラッチの逆極性のクロックでラッチして２分の１サイクル遅延させて転送する。

【００９７】［回送モード］回送モードでは、転送ラッチの内容を回送ラッチに移せばよい。但し、転送ラッチのタイミングの位相が回送ラッチのセットアップ時間内にあれば、いったん調整ラッチに移してから回送ラッチにラッチする。モジュールの接続数が少ない場合にリング１周のサイクル数を２サイクルにするため、或いは後述するように２つのリングのリング１周のサイクル数を同一にするために調整ラッチ１を設ける。回送ラッチのラッチタイミングの前後の内容が同じであることで、転送ラッチと回送ラッチのラッチタイミングの重なりのないことを確かめる。回送モードのモジュールは、１０の繰り返しパターンを送信する。回送ラッチのタイミングを規定値だけ遅延させて転送ラッチの出力をラッチする。これは回送ラッチのタイミングの後の転送ラッチの出力を見ることになる。転送ラッチの出力を転送ラッチのタイミングを規定値だけ遅延させて、回送ラッチのタイミングでラッチする。これは回送ラッチのタイミングの前の転送ラッチの出力を見ることになる。両者の信号が一致すれば、回送ラッチと転送ラッチのタイミングの重なりがないので調整ラッチを経由する必要はない。

【００９８】図４７を参照してサイクル調整回路を説明する。転送ラッチと回送ラッチのタイミングに重なりがあると、調整ラッチ使用３１１が１になる。モジュール制御ユニットは、リング制御ユニットの指示により調整ラッチ１使用３１２を１にする。転送ラッチ３１３は調整ラッチ使用３１１が１ならば調整ラッチ３１４に移す。調整ラッチ１使用３１２が１ならば調整ラッチ３１４の内容を調整ラッチ１（３１５）に移す。転送ラッチ３１３の内容を直接、調整ラッチ３１４を経由、調整ラッチ１（３１５）を経由するかを論理積３１６、論理積３１７、論理積３１８により選択して回送ラッチ３１９に移す。

【００９９】［受信線毎のスキュー補正］受信線のスキューの設計値は、動作周波数に影響を与える。選択線、アドレス線、データ線からなる受信線毎に可変遅延回路を設けスキューの自動補正を行えば動作周波数に与えるスキューの影響は補正精度だけになる。動作周波数は、ドライバ、ＬＳＩ・基板内の配線、パッケージ、コネクタ等の伝送特性で定まる。可変遅延回路は、ゲートの遅延を利用した位相調整回路と同様のものでよい。但し、反転入力、調整用の入力信号は不要である。最小の遅延量はリングの設計値で決まり２分の１サイクルまでの必要はない。

【０１００】図４８に受信線毎にスキュー補正する転送モードの概念図を示す。受信位相の時と同様に位相設定用の１０００のパターンを前のモジュールより送信する。受信線毎の可変遅延回路に遅延量を０にして最も遅い位相の受信線６１１に調整クロックの位相を合わせる。受信線毎の可変遅延回路の遅延量を一斉に増やし、受信線毎に転送ラッチで１を検出する最後の可変遅延回路の遅延量を格納する。遅延の走査完了後一斉に格納し

た遅延量に戻す。この結果、すべての受信線の転送ラッチの入力は、同じ位相になる。調整クロックの元のクロックの逆極性のクロックを位相調整回路に供給することにより２分の１サイクル遅延させ受信位相のクロックとする。

【０１０１】［全報知のパケット］パケットを各モジュールに同時に送信するために、全報知信号（１本）を設ける。全報知信号はアドレスと同時刻に１にすることにより全報知を示す。全報知信号が１の時にはアドレス信号に要求元のモジュールのアドレスを示す。全報知信号は、説明を簡潔にするために信号線として分離しているが以下の説明で明らかになるように実際にはパケットのヘッダに含められるので信号線の本数を増加させない。

【０１０２】［フロー制御］リングのレベルでパケットの送達確認をパケットによらず２本の信号線の追加で高速に行い、複雑な制御回路を回避し、スループットの低下を除去する手段を設ける。全報知のパケットに対しての送達確認の手段も設ける。リングのレベルでの再送信を減らすために、トラフィックに応じたフロー制御のアルゴリズムを与える。

【０１０３】［送達確認］図４９により送達確認の説明をする。受信モジュール４００が、受信の２サイクル後に受信不能４０２を送信モジュール４０３に報告する信号（１本）４０１を設ける。送信モジュール４０３が受信不能の報告を受けたことを報告受領後２サイクル後に各モジュールに通知する信号（１本）４０４を設ける。受信不能を報告した受信モジュール４００は、受信可能になったことを全報知のパケットで送信する。送信モジュール４０３は、２サイクル後の送信パケット（４０８）に加え、その２サイクル後の受信不能信号４０１、受信不能信号４０１を送信モジュール４０３が受信してから２サイクル後の受信不能通知信号４０４の送信権を獲得する。

【０１０４】［受信不能］送信モジュール４０３はパケット送信の２サイクル後に受信不能信号４０１を０（４０５）にする。受信モジュール４００は、パケットを受信できなかった時に受信不能信号４０１に１（４０２）を入れる。パケットを受信できた時には何もしない。送信モジュール４０３は、受信不能信号４０１が０（４０５）のままであることで送達確認をする。受信できれば何も受信不能信号４０１にしないので０のままになり全報知に対しても送達確認できる。

【０１０５】［受信不能通知］送信モジュール４０３は各モジュールに受信不能４０４を通知する。受信不能信号４０１が１（４０２）ならば、受信の２サイクル後に受信不能通知信号４０４を１にする。受信不能信号４０１が０（４０５）のままならば受信の２サイクル後に受信不能通知信号４０４を０にする。各モジュールは、特定の受信モジュールに宛てたパケットが受信不能である場合は、特定の受信モジュールが受信不能であることを知る。各モジュールは、全報知のパケットが受信不能である場合は全報知パケットが送信不能であることを知る。送信モジュールに到達する前の受信不能信号４０１が１（４０２）であることをスヌープ（４０６）できる受信モジュール４０７はそれで受信不能を知ることができるので処理に利用しても良い。

【０１０６】［受信不能原因の報告］受信不能線に加え、原因別の受信不能原因線を追加して受信不能原因を報告してより強力なフロー制御を行うことができる。同様に受信不能通知線に加え、原因別の受信不能原因通知線を追加して、より強力なフロー制御を行うことができる。

【０１０７】［再送信］特定の受信モジュールに宛てたパケットが受信不能になると各モジュールは、そのモジュールへの送信及び全報知の送信を抑止する。全報知のパケットが受信不能になると各モジュールは、全報知の送信を抑止する。受信不能を報告した受信モジュールの受信バッファの空きエントリがリングのサイクル数の２倍以上になると受信可能を全報知のパケットで送信する。各モジュールは、その受信モジュールが受信可能になったことを知る。受信不能の受信モジュールに宛てたパケットの送信を再開する。送信不能である全報知のパケットの再送信は、複数の受信モジュールが受信不能であることがあり得ることから特定の受信モジュールに宛てるパケットの後にする。全報知のパケットの再送信は、空きがリング１周のサイクル数連続するか規定の時間送信を抑止する。全報知のパケットの受信不能が増加すれば、送信を再開する迄の規定の時間を長くする。全報知のパケットの受信不能がなければ、送信を再開する迄の規定の時間を短くする。

【０１０８】［フロー制御］フロー制御は、自分のパケットの送信権の放棄により行う。次に接続されているスロットは最高の順位の送信権を得る。空きパケットが、連続して規定値以上のサイクルの間に存在しない場合は引き続き送信するパケットがあっても送信権を放棄する。規定値は、平均のビジー率が低ければ自動的に大きく設定する。バースト的なビジー状態でのビジー率の低下を防ぐ。平均のビジー率が高ければ、特定の受信モジュールに宛てるパケットを優先し、ビジー率が低ければ全報知のパケットを優先する。

【０１０９】［複数サイクルのパケット］パケットを複数サイクルで転送することにより、選択線を中心とする各モジュールの論理動作のサイクル数が増加し、またモジュール間の信号線の本数の削減によりコストの低減を図れる。２サイクルのパケットに対しては２分の１のクロックを供給しているので配信される２分の１のクロックの位相とパケットの最初のサイクルの関係を各モジュール毎に固定する。２サイクル以上の場合は、パケットのサイクル数の周期のパケットクロックを各モジュールで作成する。パケットクロックは、最初のサイクルが１

で残りのサイクルが０である。各モジュールでのパケットクロックの同期は、リング制御ユニットから同期指示を出し、回送モードのモジュールはパケットのサイクルの最初を１残りを０にする繰り返しパターンを同期パターンとして送信する。各モジュールは、特定の同期パターンを受信してパケットクロックを同期させる。送信権の制御は、選択線を使用してパケットクロックが１であるサイクルに行う。３サイクル目から選択線は、０にされる。これにより、パケットクロックの同期状態を各モジュールで監視する。後述するディレクトリ・リング或いは選択リングは、複数サイクルに分けて送信するので２サイクルのパケットであれば例えばそれぞれ１１本が６本に、２４本が１３本になる。

【０１１０】［ブロックパケットの転送］キャッシュラインの様な固定長のメッセージを送信する場合、連続的にパケットを送信することが保証されていれば、パケットへの分割転送の無駄が省けてスループットが向上し、受信モジュールでのメッセージの組立が容易になる。ブロックパケットの送信を行う時には、あらかじめブロックパケットの送信を通知して通常のパケットの抑止を行う。ブロックパケットの送信権の獲得は、ブロックの最初のサイクルで行う。ブロックパケットの送信権は、リング１周のサイクル数がブロックパケットのサイクル以下ならばブロックパケットのサイクル数の後に引き継ぐ。ブロックパケットの送信権は、リングの１周のサイクル数がブロックパケットのサイクル数以上ならばリング１周のサイクル数後に引き継ぐ。通常のパケットには、リング１周のサイクル数だけの独立した送信権がある。リングの１周のサイクル数がブロックパケットのサイクル数以上ならばその端数を除いた倍数だけの独立した送信権がある。ブロックパケットの送信をすることを予約し、それを各モジュールに通知して通常のパケットの送信を抑止してからブロックパケットの送信を行う。

【０１１１】図５０を参照して独立した送信権が２つの場合を説明する。ブロックパケットを送信する予約権を獲得するために、予約信号（１本）を設ける。ブロックパケットの予約権はサイクル毎に存在しないので、その存在を最初に明示的に示すために起動する。これは１つの信号からなるトークンを発行することを意味する。パワーオン時にアドレス０のモジュールは、予約信号を１サイクルの間１（予約１（４５０））にして予約権を１つ起動して次のサイクルから単に中継する。最初の予約信号の送信後、リング１周のサイクル数にいたる迄にブロックパケットのサイクル数があれば独立した送信権を必要数（予約２（４５１））起動する。ブロックパケットを送信したいモジュール（４５２）は、予約信号に０を送信し続ける。ブロックパケットを送信したいモジュール（４５２）は１（予約１（４５０））の予約信号を受信すると予約権を獲得し、２サイクル後からは予約信号には０を送信せず単に中継する。ブロックパケットの

サイクル数は２以上なので次の予約信号（予約２（４５１））には影響がない。ブロックパケットの予約権を通知するため予約通知線（１本）を設ける。予約権を獲得の２サイクル後から予約通知線に１サイクルの間１（予約通知１（４５３））を乗せ、リング１周後のブロックパケットのサイクル数分（この場合４）の送信権の予約を各モジュールに通知する。予約権を獲得したモジュール（４５２）はリング１周のサイクル後予約通知線に１サイクルの間０（予約３（４５４））を送信する。ブロックパケットの予約権を獲得したモジュール（４５２）は、予約通知の２サイクル後に選択線を１にして送信要求（送信要求１（４５５））を行い送信権を獲得する。ブロックパケットのサイクル数の間も同様に送信権を獲得できる。

【０１１２】次の予約権はブロックパケットのサイクル数がリング１周のサイクル数より大きいならば予約した後のブロックパケットのサイクル数の後になる。この場合はブロックパケットのサイクル数がリング１周のサイクル数以下なので、次の予約権はリング１周のサイクル後になる。ブロックパケットを引き続いて送信しないならば、予約線を１サイクルの間１（予約３（４５４））にして予約権を放棄する。次のサイクルからは１を送信せず単に中継する。引き続いてブロックパケットを送信したい場合は、何もせずに予約権を獲得する。各モジュールは、予約通知の受信の２サイクル後からブロックパケットのサイクル数の間通常パケットの送信要求を抑止する。

【０１１３】ブロックパケットの送達確認は、ブロックパケット受信不能信号、ブロックパケット受信不能通知信号を設けて行う。ブロックパケットが受信不能ならば、最初のパケットの２サイクル後にブロックパケット受信不能信号を１にして報告する。送信モジュールは、受信不能信号２サイクル後にブロックパケット受信不能通知信号を１にしてブロックパケットの送信不能を通知する。送信モジュールはその受信モジュールに対するブロックパケットの送信を抑止する。受信不能を報告したモジュールは、全報知パケットによりブロックパケットが受信可能になったことを通知する。通常のパケットの送達確認と同様に、原因別の受信不能報告及び通知の機能を拡張できる。通常のパケットとしての送達確認は、受信不能信号、受信不能通知信号を使用する。平均ビジー率が定められた値より大きければ連続して送信権を獲得しない。これにより送信権を獲得した次のアドレスのモジュールが優先的に送信権を獲得できる。予約権を保有しているモジュールは明示的でないのでアドレス０のモジュールで監視を行う。アドレス０のモジュールはフロー制御と予約信号の正常性の監視のため、予約信号がブロックパケットのサイクル数或いはリング１周のサイクル数のどちらかの周期以上で受信すること且つあらかじめ定められた一定時間以内に受信することを確認す

る。定められた数のブロックパケットを連続して送信することは、予約線、予約通知線と同じ働きをする長ブロック予約線、長ブロック予約通知線を追加することで実現できる。長ブロックパケットを複数個纏めた超パケットを送信するには、長パケット予約線に相当する超パケット予約線、長パケット予約通知線に相当する超パケット予約通知線¥を設ける。同様に、送信権獲得のための信号を追加することにより階層的な構造の異なる長さを持つ複数種類のブロックパケットを送信できる。図５１にリング１周のサイクルより大きなブロックパケットの転送の例を示す。

【０１１４】［ディレクトリ・リング］ディレクトリ・リングにより、モジュールのキャッシュにデータを保有していることを送信モジュールに報告し、送信モジュールがそれを各モジュールに通知することができる。送信モジュールは送信権を獲得して報告を要求し、報告に基づいてそれを通知する。図５２にキャッシュにある報告がある場合、図５３にキャッシュにない場合、図５４にその報告の要求では確定せず、引き続き報告要求をする場合を示す。

【０１１５】ディレクトリ・リングの送信権の制御を行う要求線５００、読み出しパケットに対応する要求番号を乗せる要求番号信号５０１（６本）、キャッシュラインの存在の有無を確認中であることを報告する未確定信号５０２（１本）、キャッシュラインがあることを報告する保有信号５０３（１本）、その結果を通知する確定信号５０４（１本）、保有通知信号５０５（１本）を設ける。信号線を要約すれば、１１本の追加になり、２サイクルでパケットを転送する場合には６本の追加になる。先に述べるディレクトリのローカルメモリのアクセス待ちの様な未確定の状態が完了する迄に多くの時間を要することを通知する長期未確定信号（１本）を設けることもできる。再要求を行う時間を待たせて、無駄な要求を減らすことができる。要求線５００により２サイクル後の要求番号信号５０１、要求線５００より４サイクル後、要求番号信号５０１から２サイクル後の未確定信号５０２及び保有信号５０３、要求線５００よりリング１周のサイクルに６サイクルを加えたサイクル後の確定信号５０４及び保有通知信号５０５の送信権の制御を行う。報告を要求するモジュールは、送信権を獲得（５０６）して２サイクル後にその要求番号を要求番号信号（５０７）に乗せる。引き続いて、未確定信号５０２、保有信号５０３を０（５０８）にする。キャッシュ状態を報告するためモジュールは、要求線が１（５０６）になっている２サイクル後の要求番号（５０７）を読む。モジュールは、その要求番号（５０７）の示すラインがキャッシュになければ、何もしないでリング１周のサイクルの２サイクル後の確定信号（５０４）を待つ。モジュールは、その要求番号（５０７）の示すラインがキャッシュにあれば、保有信号（５０３）を１（５０９）にし処理を完了する。

【０１１６】図５２の未確定信号（５０２）のｘ（５１０）は、０あるいは１を示す。図５２、図５３、図５４において、リング３周の０サイクル目の受（５１１）は送信モジュールが受信することを示す。モジュールは、その要求番号（５０１）の示すキャッシュラインの存在の有無の確認中は、未確定信号（５０２）に１（図５４の５１２）を入れ、リング１周のサイクルの２サイクル後の確定信号（図５４の５１３）を待つ。要求モジュールは、リングの１周のサイクル後の未確定信号（５０２）と保有信号（５０３）を受信する。要求モジュールは、読み出した保有信号（５０３）が１（図５２の５０９）か、未確定信号（５０２）が０（図５３の５１４）ならば処理が終わったので、２サイクル後に確定信号（５１５）、保有通知信号（図５２の５１６）を１にして処理を完了する。要求モジュールは保有信号（５０３）が０（図５４の５１７）で未確定信号（５０２）が１（図５４の５１２）ならば再び要求信号５００を１（図５４の５１８）にして送信権を獲得して報告を要求する。モジュールは、未確定信号（５０２）、保有信号（５０３）に応答後リング１周のサイクルの２サイクル後の確定信号５１５が１（図５２、図５３の５１５）になったら保有信号（５０３）、未確定信号（５０２）の処理を中止或いは完了する。確定信号（５０４）が０（５１３）なら、要求モジュールからの再要求に備える。モジュールは、保有信号（５０３）が１（５０９）であれば最終状態であることをスヌープできる。メモリモジュールは、読み出しデータがあり、保有通知信号（５１６）が１（図５２の５１６）であれば送信を抑止する。メモリモジュールは、読み出しデータがあり保有通知信号（５１６）が０（図５３５１９）であれば送信する。

【０１１７】［論理的順序付け］キャッシュの所有権を決めるため、全報知のパケットはその論理的な順序付けを行い、各モジュールが認識する必要がある。全報知のパケットは、リングを１周する間送信権が獲得されているので必ずアドレス０のモジュールを通過する。アドレス０のモジュールはパケットが通過する順に論理的順序を定めることができる。受信モジュールはアドレス０のモジュールの論理的順序に合わせる。受信モジュールとアドレス０のモジュールとの間のモジュールが送信モジュールがあると、アドレス０のモジュールを通過して、受信モジュールを通過し送信モジュールに戻る。パケットの受信時刻で論理的順序を定める。アドレス０のモジュールと受信モジュールとの間に送信モジュールがあると、受信モジュールを通過してアドレス０のモジュールを通過し送信モジュールに戻る。アドレス０のモジュールを通過する順序は、仮想的に送信モジュールから受信モジュールまでパケットを送信する場合の順序になる。パケットの受信時刻にリング１周のサイクル数を加えた

時刻で論理的順序を定める。送信モジュールのアドレスは、リングの宛先のアドレス線に示されている。図５５は、アドレス０のモジュール５５０を通過するパケットＡ、Ｂ、Ｃ、Ｄ、Ｅ、Ｆが各モジュールにおいて受信される順序（５５１）を示す。受信モジュールは、自アドレスと宛先アドレスを比較して論理的順序を定める時刻を（５５３）になるように定める。パケットＡ（５５２）はアドレス２のモジュールが送信することを示す。図５６は要求元アドレスと自アドレスを比較して自アドレスが大きければ論理的順序づけの時刻は現時刻にリング１周のサイクル数を加えたものになることを示す。

【０１１８】［選択リング］次に選択リングについて説明する。基本リングの選択線による送信権獲得の代わりに選択リングを用いてアドレス０のモジュールで集中的に送信権の獲得を行うことができる。集中的に送信権獲得を行うために１サイクルの間に２つのパケットの転送を行うことができるのでスループットを向上できる。また、アドレス０のモジュールで処理するので、論理的順序付けは送信権の獲得順でよく、各モジュールで特別の処理はいらない。選択線は分散的送信権獲得を行うのに対し、選択リングは集中的送信権獲得を行う。要求モジュールからアドレス０のモジュールまでの転送時間、アドレス０のモジュールでの送信権処理時間、アドレス０のモジュールから要求モジュールまでの転送時間がかかるのでレイテンシは増加する。信号線の数も増加する。選択線を用いた送信権はリング１周の間、獲得される。全報知のパケット以外は要求元から宛先アドレス迄リングを占有する。宛先アドレスから要求元までは、別のパケットを送信できる可能性がある。選択線を用いた送信権制御は各サイクルの送信権をリング１周のサイクルで処理する。送信権制御を集中するとリングの利用可能の領域に別のパケットを送信できる。選択リングを設け、集中送信権制御をアドレス０のモジュールで行う。集中送信権制御では、１つの送信権の制御にリングの１周のサイクルを使っているのを２つのパケットの送信権制御を１サイクルで処理する。各モジュールからアドレス０のモジュールには各モジュールからの送信要求を報告し、アドレス０のモジュールから獲得した送信権を通知する。キャッシュ制御を行うマルチプロセッサシステムでは、読み出しとキャッシュのライトバックは、同時に発生することが多く、また通常のパケットは全報知である場合が多い。全報知でない２つのブロックパケットをリング１周に並行に転送し通常のパケットはアドレス０から最終アドレスのモジュールの間に１つだけ転送する方式を述べる。論理的順序づけは送信権の順に行われる。ブロックパケットの予約線とそれに付随する信号線は必要でない。各モジュールからアドレス０のモジュールへの送信要求の報告信号は、選択線（１本）及び送信権報告線（１０本）である。送信権報告線（１０本）には、送信要求報告の時には送信要求（１本）、要求元アドレス（３本）、宛先アドレス（３本）、ブロックパケット（１本）、全報知（１本）を乗せる。送信権報告線（１０本）には、ビジー報告の時には送信要求を０にしてビジー報告（１本）、ビジー報告アドレス（３本）、ブロックビジー（１本）、ビジー（１本）を乗せる。送信権報告線には、ビジー解除報告の時には送信要求を０にして、ビジー解除報告（１本）、ビジー解除報告アドレス（３本）、ブロックビジー（１本）、ビジー（１本）を乗せる。アドレス０のモジュールから各モジュールへの送信要求通知信号は、送信権有効（１本）、送信権通知線（１２本）である。送信権通知線は、送信要求通知の時にはそのサイクルで獲得した送信権０と送信権１の２つまで通知できる。送信権０は送信要求通知０（１本）、要求元０アドレス（３本）、ブロックパケット０（１本）、全報知０（１本）を乗せる。送信権１は送信要求通知１（１本）、要求元１アドレス（３本）、ブロックパケット１（１本）、全報知１（１本）を乗せる。送信権通知線は、ビジー通知の時には送信要求通知０を０にして１つのモジュールに対するビジー通知（１本）、ビジー通知アドレス（３本）、ブロックビジー（１本）、ビジー（１本）を乗せる。送信権通知線には、ビジー解除通知の時には送信要求通知０を０にして、１つのモジュールに対するビジー解除通知（１本）、ビジー解除通知アドレス（３本）、ブロックビジー（３本）、ビジー（１本）を乗せる。必要な信号線の本数を要約すると報告に１１本、通知に１３本の合計２４本である。２サイクルでパケットを転送する場合には１３本が追加になる。フロー制御における送達確認は、１サイクルに２つ並行転送することから受信不能（１本）、受信不能通知（１本）をそれぞれ２本に拡張する。各モジュールからアドレス０のモジュールへパケットの送達確認のため送信権０、送信権１に対応して受信不能報告０（１本）、受信不能報告１（１本）を設ける。各モジュールからアドレス０のモジュールへパケットの送達確認のため送信権０、送信権１に対応して同様に受信不能通知０（１本）、受信不能通知１（１本）を設ける。各モジュールは、送信要求数と処理完了の応答数の差から未処理送信要求数を監視してアドレス０の送信要求バッファが満杯にならぬように送信要求を制御する。同一の要求元からの全報知のパケット、ブロックパケットは要求の順に応答するので送信要求を要求番号まで特定する必要はない。

【０１１９】［送信権獲得処理］通常のパケットの集中送信権制御の処理は基本リングで行われている処理をアドレス０のモジュール内の論理回路だけで実現する。アドレス０のモジュールでの集中送信権制御を以下に示す。サイクルのあるアドレス（起点）以降からそのサイクルの最後のアドレス（終点）の送信権を定める。選択する送信要求がなければ、アドレス０をリング１周のサイクル後の送信権処理の起点としてそのサイクルの処理

を完了する。選択する送信要求の宛先がリングのアドレスの終点より先にあれば、宛先のアドレスをリング１周のサイクル後の送信権処理の起点として処理を完了する。その宛先が終点より前にあれば、アドレス０をリング１周のサイクル後の送信権処理のの起点としてそのサイクルの処理を完了する。

【０１２０】［高速の送信権獲得回路］送信権獲得のクリティカルパスは２入力の論理和回路の従属接続にある。図５７に各モジュール毎の送信権獲得論理を示す。図５８に図５７の送信権獲得論理を８個従属接続して各モジュールの送信要求から送信権を獲得する論理を示す。図５９に接続数の多い場合に図５８の回路を並列に４個接続して送信権を獲得する論理を示す。

【０１２１】［並列転送ブロックの選別］ブロックパケットに対してはリング１周の間に最大２つのパケットの送信権を処理する。最初のブロックパケットに対してそのサイクルのあるアドレス（起点）以降からそのサイクルの最後のアドレス（終点）の送信権を通常のパケットと同様に定める。送信権を獲得したブロックパケットがあれば、そのブロックの宛先アドレスから要求元迄の間のリングの領域を使用するブロックパケットを並列転送可能なブロックパケットを選別する。図６０に送信要求の中で並列転送可能なパケットブロックの選別回路を示す。

【０１２２】［並列転送ブロックの選択］並列転送可能なパケットブロックのうちでアドレスの最も小さいブロックパケットが送信権を獲得する。図６１は各モジュールでの並列転送可能ブロックから並列転送ブロックを１つ選択する回路を示す。図５９の並列転送可能ブロックの選別回路を８個従属接続する。ブロックパケットの送信権獲得のクリティカルパスは２入力の論理和回路の従属接続にある。標準的なゲートによらず専用の回路を用いて良い。

【０１２３】［並列転送ブロックの通知］２つのブロックパケットが同一サイクルにあれば、送信権獲得を同時にリングを用いて通知する。２つ目のパケットが次のサイクルにあれば、次のサイクルで通知する。ブロックパケットは少なくとも2サイクル続くので、その送信権処理に２サイクル使用する。

【０１２４】［フロー制御］フロー制御として、平均ビジー率の低い場合は全報知を優先し、平均ビジー率の高い場合はブロックパケット、全報知でないパケットを優先する。平均ビジー率が高い場合は同のモジュールで定められた数を越えて連続して送信権を獲得できない。

【０１２５】［階層接続］リング接続したプロセッサシステムの能力を向上させるために階層接続する。プロセッサシステムを副リングとして、副リング間を主リングで階層的に接続する。副リング内で閉じる転送は主リングに影響しない。全体のレイテンシは主リングと副リングのレイテンシの和でレイテンシが決まる。主リングからのスループット要求を満たすために副リングのモジュールの接続数を低く押える必要のある時には、副リングをケーブル等で短絡してレイテンシを改善できる。副リングは、多くの場合１つのバックプレーンに収容される。バックプレーン間を接続する主リングは、リングの一つの特徴であるケーブルで接続できるので容易に実装できる。規定の長さを超えて接続する場合は、延長のためのモジュールを追加する。図６２に階層接続の接続方法を示す。主リングと副リングの接続機構は、副リングのモジュールのアドレス０を持ち、副リングへの送信権獲得の通知と主リングへの送信要求を同時に行うことができる。接続機構はパケットのヘッダより主リング内のアドレス或いは副リング内のアドレスを生成する。なお、副リングは通常のバス構造でも可能である。メッセージのセマフォア、アドレスマップレジスタ、ＤＭＡレジスタ等ＩＯアドレス空間を共有する場合は、キャッシュしないので宛先が１つに定められパケットの転送は１対１になる。分散共有メモリのように全報知のパケットが生じる場合は、各副リングに必要なスループットは接続する副リングの総数に直線的である。

【０１２６】図６３に階層接続のキャッシュ可能な共有メモリを有する接続機構を示す。副リング７００内で閉じるパケットは、主リング７０１には転送しない。当該副リング７００に宛てた主リング７０１のパケットは、副リング７００に転送する。全報知のパケットは、すべての副リング、主リングに転送する。レイテンシを短くできるがリングに必要なスループットは増加する。接続機構にＳＣＩに定められたディレクトリ、プロトコルを用いれば全報知パケットをなくすことができるが、レイテンシは増加し、ディレクトリは大きくなる。リングの全報知能力を活かして接続機構にキャッシュ・コヒーレンスのための全報知パケットを削減し、レイテンシの増加を抑制し、適当な大きさのディレクトリを持つことができる。副リング７００に属するモジュールがその副リングに属するローカル・メモリ７０３にアクセスする場合、他の副リング７０４のキャッシュに写しがなければ主リング７０１にそのパケットを転送しないでよい。副リング７００に属するローカルメモリ７０３の写しを他の副リング７０４のキャッシュに持っていないことを示すローカル・ディレクトリ７０５を持てば、主リング７０１へのパケットの転送を削減できる。ある副リング７００に属するモジュールが、他の副リング７０６に属するローカル・メモリ７０７をアクセスする場合、アクセスされたローカルメモリ７０７を持つ他の副リング７０６がその副リング７０４には写しを持っていないことを通知すれば、その副リング７０４内にパケットを転送しないでよい。他の副リング７０６に属するローカルメモリ７０７が、その副リング７０４のキャッシュに写しのある状態を示すリモート・ディレクトリ７０６を持てば、その副リング７０４内に転送するパケットを削減で

きる。宛先の副リングは３６ビットの内の最初の４ビットを用いて行われる。キャッシュ可能なメモリ領域であり実装されていることを確認してからディレクトリをアクセスする。

【０１２７】［ローカル・ディレクトリ］メモリアドレスは３６ビット（６４ＧＢ）で、副リングに実装するローカルメモリのアドレスは３２ビット（４ＧＢ）、ラインの大きさは６ビット（６４Ｂ）とする。ローカル・ディレクトリのエントリ数は、３２ビットから６ビットを引いた２６ビット（６４Ｍエントリ）である。

【０１２８】図６４にアドレスの関係を示す。ローカル・ディレクトリのエントリには、写しビットと変更ビットの２ビットを持つ。写しビットは、他の副リングのキャッシュに写しがあれば１、なければ０を格納する。変更ビットは、他の副リングのキャッシュに写しがあり変更されている可能性があれば１、なければ０を格納する。ローカル・ディレクトリは、ローカルメモリの２５６分の１の容量である。

【０１２９】図６５にローカル・ディレクトリを示す。メモリアドレスを４０ビット（１０２４ＧＢ）に、副リングに実装するローカルメモリのアドレスを３６ビット（６４ＧＢ）に拡張する場合は、ローカル・ディレクトリのエントリ数は３６ビットから６ビットを引いた３０ビット（１０２４Ｍエントリ）になる。大容量であるので、ローカル・ディレクトリ・テーブルとしてローカルメモリ内に格納してローカル・ディレクトリにはその写しを持つ。ローカル・ディレクトリの写しは、ラインの大きさと同じ６４Ｂ単位の大きさにする。６４Ｂ（５１２ビット）には２５６（８ビット）のエントリを収容できる。ローカル・ディレクトリには、２６ビットから８ビットを引いた１８ビット（２５６Ｋエントリ）の写しを格納できる。ローカルメモリのアドレスの３６ビットから８ビットと６ビットを引いた２２ビットを１８ビットに変更するためにアドレス変換テーブルを持つ。

【０１３０】図６６にアドレスを拡張する場合を示す。アドレス変換テーブルは、２５６Ｋエントリ（１８ビット）を持ち、各エントリには２２ビットのアドレスを持つ。アドレス変換テーブルの１つのエントリは、２２ビットから１８ビットを引いた４ビットに相当する１６単位の内の１つの写しを示す。

【０１３１】図６７にアドレス変換テーブルを示す。ローカルディレクトリは、ローカルメモリのすべての範囲のキャッシュの状態を格納しているのでディレクトリリングは必要ない。アドレス変換テーブルのエントリ数も、１８ビットのアドレスをハッシュすることにより減らせる。

【０１３２】［リモート・ディレクトリ］各エントリには、副リングに属するローカルメモリの写しがその他の副リングにあれば１、なければ０を主リング内の副リングのアドレス０からアドレス７まで順に格納する。この８ビットを写しベクトルと呼ぶ。主リング内の副リング（アドレス）を３１に拡張するときには、各ビットをグループ単位に割り当てる。グループ内はＳＭＰでグループ間はクラスタの場合等に効果がある。例えば第１ビット目は、アドレス０－３のグループの様に割り当てる。４つのアドレスの中の１つにしか写しがなくても、４つのアドレスに対応する副リングすべてに写しがあるとして取り扱う。リモートディレクトリのエントリの数は、再利用する頻度によって決められる。２２ビット（４Ｍエントリ）にする場合は、そのエントリを特定するためにローカルアドレスの３２ビットから２２ビットと６ビットを引いた４ビットを各エントリに格納する。ローカルアドレスを３６ビットに拡張するには、エントリを特定するために３６ビットから２２ビットと６ビットを引いた８ビットを格納する。図６８にリモート・ディレクトリを示す。

【０１３３】［写しベクトルの通知］副リングに属するローカルメモリにアクセスする全報知のパケットを主リングに転送する時には写しベクトルをヘッダに格納して転送する。副リングに属するローカルメモリにアクセスする全報知のパケットを受け取った時には写しベクトルをパケットを用いて通知する。写しベクトルは自身のアドレス位置には写しビットはいらないのでその位置が１であれば有効であることを示す。

【０１３４】［ローカルメモリの種別］ローカル・ディレクトリによって、ローカルメモリは写し無し（Ｎ）、写し有り（Ｃ）、変更有り（Ｍ）に分けられる。リモート・ディレクトリによって、ローカルメモリの写し有り（Ｃ）の一部は写しベクトル有り（ＣＨ）、変更有り（Ｍ）の一部は写しベクトル有り（ＭＨ）に分けられる。図６９にディレクトリにより分けられるローカルメモリの種別を示す。

【０１３５】［ローカルメモリアクセス］副リング７０５に属するモジュールがローカルメモリ７０３を写すパケットを送信して、接続機構７０８がこれを受信したら、Ｎ、Ｃなら主リングへ転送しない。Ｍなら変更した内容をキャッシュより書き戻す必要がある。このため、写しベクトルをヘッダに格納して主リング７０１に全報知のパケットを送信する。主リング７０１の他の副リング７０４は、写しベクトルが有効で且つ０であれば副リング７０４内にパケットを転送しない。写しベクトルが有効で且つ０か無効であれば副リング７０４内にパケットを転送する。他の副リングから書き戻される間、接続機構において同一アドレスに対するアクセスは待たされる。書き戻されるとＣＨとなり処理が引き継がれる。副リング７００に属するモジュールがローカルメモリ７０３を変更するパケットを送信して、接続機構７０８がこれを受信した場合、Ｎなら主リングに転送しない。Ｃならキャッシュの写しを廃棄する必要がある。Ｍならキャッシュの内容を書き戻した上で廃棄する必要がある。

Ｃ、Ｍなら写しベクトルをヘッダに格納して主リング７０１に全報知のパケットを送信する。主リング７０１の他の副リング７０４は、写しベクトルが有効で且つ０であれば副リング７０４内にパケットを転送しない。写しベクトルが有効で且つ０か無効であれば副リング７０４内にパケットを転送する。Ｍなら書き戻されるのを待ちＮに、ＣならそのままＮになる。以上で明らかな様に、ローカルメモリで閉じる処理は他の副リングの影響を受けない。

【０１３６】［リモートメモリ・アクセス］副リング７００に属するモジュールが他の副リング７０６に属するローカルメモリ７０７（リモートメモリ）を写すパケットを送信して接続機構７０８がこれを受信したら、主リング７０１に転送する。ローカルメモリ７０７を有する副リング７０６は、写しベクトルをパケットにより通知する。主リング７０１の他の副リング７０４は、写しベクトルの受信してから受信してあるパケットの処理を行う。写しベクトルが有効で且つ０であれば副リング７０４内にパケットを転送しない。写しベクトルが有効で且つ０か無効であれば副リング７０４内にパケットを転送する。ローカルメモリ７０７の状態はＣＨになる。副リング７００に属するモジュールが他の副リング７０６に属するローカルメモリ７０７（リモートメモリ）を変更するパケットを送信して接続機構７０８がこれを受信したら、主リング７０１に転送する。ローカルメモリ７０７を有する副リング７０６の接続機構７０９は、パケットを受信すれば副リング７０６内に転送する。ローカルメモリ７０７を有する副リング７０６の接続機構７０９は、Ｎならローカルメモリ７０７をアクセスし、また他の副リング７０４でパケットの処理をしないで良いことをパケットにより通知する。Ｃならローカルメモリ７０７をアクセスし、また他の副リング７０４のキャッシュに写しがあれば廃棄することを通知する。Ｍならば、書き戻しを行ってから廃棄することを通知する。これらのため、キャッシュの写しベクトルをパケットにより通知する。主リング７０１の他の副リング７０４は、パケットから写しベクトルを受信してから受信してあるパケットを処理する。写しベクトルが有効で且つ０であれば副リング７０４内にパケットを転送しない。写しベクトルが有効で且つ０か無効であれば、副リング７０４内にパケットを転送する。ローカルメモリの種別はＭＨになる。以上で明らかな様にリモートメモリにアクセスする処理は、レイテンシを遅くすることなくまた必要のない副リングへのアクセスを行わない。

【０１３７】［双方向リング］モジュール間は、１対１の１方向の伝送を行っている。双方向伝送回路を用いて２つの独立する双方向リングを構成することができる。双方向リングはリング接続の配線をそのままにして性能の限界を２倍にする増設性を与える。２つのリングの同期ができないと言う制約があり、伝送の方向性の弁別回路の弁別精度、周波数特性の制約から動作周波数の限界は低くなる。また弁別回路そのもの及び弁別回路が発生する雑音のために中継モードのモジュールの数の制限が厳しくなり、転送モードのモジュールの割合が増えてレイテンシが長くなる。全報知の通常パケットは１つのリングを利用して論理的順序性の保障を行い、順序性の制約のない通常のパケット、ブロックパケットを２つのリングに分散させる。２つの並列接続リングの構成に比較すれば、性能は若干低くなる。図７０に弁別回路の例を示す。図７１に双方向リングのモジュール間の接続を示す。

【０１３８】［構成変更］リング制御ユニットは、各モジュールへのヘルスチェックのフレームを送信して、応答のないことで構成変更が行われていることを検知する。リング制御ユニットは、各モジュールへのヘルスチェックを行い、再応答により構成変更が完了したことを知り、イニシャライズを行う。構成変更によって一般にいくつかのパケットが失われるので、適当なリスタート・ポイントまで遡る必要がある。リングの構成変更の際に、計画停止してからリングの動作を停止しイニシャライズを行った後に直ちに再開する方法を図７２に示す。モジュールの引き抜きを早期に知るために、モジュールの引抜レバーの引き起こしを引抜センサで検知してモジュール制御ユニットを通じてリング制御ユニットに報告する。リング制御ユニットは、引き抜きの報告により各モジュールに対して新たな動作要求を開始せず、それ以前の動作要求への応答は行いしかる後ハードウェア動作を正常停止することを指示する。モジュールの引抜に連動して短絡スイッチは短絡する。リング制御モジュールは、ヘルスチェックによりモジュールの引抜の完了を自動的に確認する。モジュールの引抜中もリングのクロックは供給されるので引抜モジュール以外のモジュールの正常停止状態は継続する。モジュールの引抜の報告よりアドレスを再設定し、スキュー補正の必要なモジュールに対してイニシャライズを行い、回送モードのモジュールでリングの１周のサイクル数を決める。リング制御ユニットは、ハードウェア停止状態を解除し、新たな構成情報を伝える。モジュールの挿入に際してもリングの動作は停止する。モジュールの挿入を早期に知るために挿入信号（１本）を設ける。挿入信号線はそのモジュールのガイドから始まり、短絡コネクタがあればこれを経由して次のモジュールに接続して終わる。挿入信号は、モジュール間の接続を行うのでモジュールは転送しない。モジュールのガイドを通過することを挿入センサで検知して、挿入信号を一時的に短絡する。各モジュールは、挿入信号が短絡されるかを常時監視していて、短絡されれば前のモジュールが挿入されるはずであるということをリング制御ユニットに報告する。リング制御ユニットは、各モジュールに対して新たな動作要求を開始せず、以前の動作要求への応答を行いしかる後ハードウェア動

作を正常停止することを指示する。モジュールの挿入に連動して、短絡スイッチを機械的にオフにする。リング制御ユニットは、ヘルスチェックの応答により挿入の完了とそのモジュールを特定する。モジュールの挿入中も、リングのクロックは供給されるので挿入モジュール以外のモジュールの正常停止状態は継続する。挿入の報告よりアドレスを再設定しスキュー補正の必要なモジュールに対してイニシャライズを行い、回送モードのモジュールでリングの１周のサイクル数を決める。リング制御ユニットは、ハードウェア停止状態を解除し新たな構成情報を伝える。２つのリングを接続して、通常状態では同期、並行動作を行い、故障時は縮退運転を行うことができる。２つのリングの同期のために回送モードにおいて１サイクルの遅延ラッチを設けリングの１周のサイクル数を一致させ、リング制御ユニットの動作を同期させる。階層接続の主リングを複数持てば拡張性は若干失われるがスループット、信頼性が向上する。共通の基板によらず、各スロット間をケーブルで接続することができる。

【０１３９】［インターフェイス信号線の具体例］インターフェイス信号の具体例を図７３に示す。この例では、メモリにディレクトリを持つのでディレクトリ・リングはない。また、選択リングは採用していない。

【０１４０】［イニシャライズ］各モジュールのアドレス、回送モード、転送モード、中継モード、位相調整回路、サイクル調整回路、可変遅延回路、リングを一周するサイクル数、複数サイクルのパケットのクロックの同期は、イニシャライズ時に設定する。リング制御ユニットは、リング全体、モジュール間の相互関係、処理の同期等を行うため各モジュールのモジュール制御ユニットに指示を出す。モジュール制御ユニットは指示に基づき可変遅延回路の設定などの処理を行う。

【０１４１】［リング制御ユニット］リング制御ユニットは、リング制御のイニシャライズを制御、統括するとともに、クロックを各モジュールにリング状に配信する。リング制御ユニットはクロックが配信されていることを確認するために受信する。各モジュールは、リング制御ユニットからの指示を基に処理し、リング制御ユニットに状態を応答する。リング制御ユニットは、フレームを用いて指示を伝達する。指示は、ステップバイステップで行う。フレームの内容は、コマンドバイト、アドレスバイト、２バイトのギャップ、２バイトのデータである。コマンドはリング制御ユニットの指示、応答要求である。アドレスは各モジュールあるいは全報知を指定する。２バイトのギャップは、モジュール制御ユニットに応答の処理する時間を与える。データは、各モジュールへの指示或いは各モジュールからの応答である。指示のフレームではアドレスで指定されたモジュールはフレームのデータを受信する。応答要求のフレームでは、アドレスモジュールで指定されたモジュールはそのフレームのデータ部に内容を入れて応答する。リング制御ユニットと各モジュールはリング状に接続する。制御リングは、各モジュールとの独立性を高めるためいわゆるリングバスの構造を取っている。フレームは、ビット、バイトからなる。ビット同期は、制御クロックで行う。制御クロックの位相はフレームのビットの中央で立ち上がる。バイトの同期は、同期パターン“００００００００００００ＦＦ”で行う。フレームの同期は、“ＦＥｘｘｘｘｘｘｘ”で行う。フレーム間は“００”を埋める。

【０１４２】［パワーオン・モード］リング制御ユニットは、全報知で各モジュールにパワーオン・モードを指示する。パワーオンモードでは回送モードと転送モードのみを使用して動作の変動に対処する。

【０１４３】［モジュールアドレスの設定］リング制御ユニットは、各モジュールに全報知でアドレス設定準備の指示をだす。各モジュールは、あらかじめ挿入するスロット番号をもっている。アドレス設定準備の指示を出されると各モジュールはアドレス部に対してスロット番号で動作する。リング制御ユニットは、アドレス部にスロット番号を入れてモジュールの種類、論理アドレス等の応答を要求する。指定されたモジュールは、データ部にプロセッサ、ＩＯ、メモリ、プロセッササブシステム等の種類を応答する。リング制御ユニットは、アドレス部にスロット番号を入れてデータ部にモジュールのアドレスを入れアドレス設定を指示する。リング制御ユニットは、各モジュールに全報知でアドレス設定完了の指示を出す。アドレス設定完了の指示を出されると各モジュールはアドレス部に対して設定されたアドレスで動作する。リング制御ユニットは、回送モード、転送モード、中継モードを設定する。

【０１４４】［受信線毎のスキュー補正］受信線毎のスキュー補正を行う場合は、リング制御ユニットは回送モード、転送モードのモジュールより１０００の繰り返しパターンの送信を指示する。引き続くモジュールに対してスキュー補正を基本動作に分解して指示する。最初に調整クロックの位相を初期設定する。可変遅延回路の遅延量の０設定、位相調整回路の調整開始の指示、受信線のすべての０の検出指示、位相調整回路の遅延量の１ステップ毎の走査、０検出の結果のステップ毎の報告により調整クロックは初期設定される。次に最大の遅延の受信線に遅延クロックの位相を合わせる。受信線のすべての１検出指示、位相調整回路の遅延量の１ステップ毎の走査、１検出結果のステップ毎の報告により位相調整回路の位相が最大の遅延の受信線に合わせられる。次に可変遅延回路の遅延量の１ステップ毎の走査により１検出の最後の遅延量が格納される。遅延量の走査の完了後、格納された遅延量を戻してスキュー補正が完了する。この場合は、遅延クロックの逆位相のクロックの供給により受信位相も同時に決定する。この動作をリングの全モ

ジュールについて行い、スキュー補正の完了を指示して遅延クロックを書くモジュールのリング制御部以外にも供給する。

【０１４５】［受信位相の決定］前のモジュールに１０００の設定パターンの送信を指示する。最初に最小の遅延の受信線に、遅延クロックを合わせこれを前縁とする。次に最大の遅延の受信線に、遅延クロックを合わせて後縁とする。引き続き受信線の中間に遅延クロックを合わせて、その逆位相のクロックを供給して受信位相が決定する。モジュール制御ユニットは、基本動作に分解してモジュール制御ユニットに指示する。

【０１４６】［サイクル調整回路］回送モードのモジュールに１０の繰り返しパターンの送信を指示する。回送モードのモジュールに、サイクル調整回路に内蔵しているクロックの規定時間前とクロックの規定時間後の入力の検出回路により、それぞれを値を応答させる。２つの値が同じであれば、サイクル調整ラッチを経由せずに回送ラッチにラッチするように指示する。２つの値が同じでなければ、サイクル調整ラッチを経由して回送ラッチにラッチするように指示する。

【０１４７】［サイクル数の設定］リング制御モジュールは、回送モードのモジュールに１つの１と引き続く指定する数の０の繰り返しパターンの送信を指示する。リング制御モジュールは回送モードのモジュールにリング１周のサイクル数を１を受信する間隔で行うように指示して、その報告を要求する。リング制御ユニットは、全報知で各モジュールにリング１周のサイクル数を知らせる。

【０１４８】［複数サイクルの位相の設定］全報知で複数サイクルの位相の設定の指示を出す。サイクル数に合わせた１に続く０の繰り返しパターンを回送モードより送信する。全報知で設定の完了を指示する。

【０１４９】［動作モードへの移行］リング制御ユニットは、パワーオン直後のスキュー量が安定状態になる規定の時間後にパワーオンモードを解除して動作モードに移行する。リングの各モジュールに対して回送モード、転送モード、中継モード、受信線毎のスキュー補正、受信位相、サイクル調整回路、サイクル数を再度決定する。

【０１５０】［誤動作検出］制御リングは、ヘルスチェックにより各モジュールを監視する。基本リング内の誤動作の検出は、データ線についてはエラー検出コード、選択線については送信権放棄が定期的に行われているかの監視で行う。リングの回復不能或いは規定回数以上の誤動作時はモジュール制御ユニットによりリング制御インターフェースを通じてリング制御ユニットに知らされる。リング制御ユニットよりリングの停止を指示してイニシャライズを行う。リング制御ユニットは各モジュールの固有の障害の報告をヘルスチェックにより受けて、モジュール制御ユニットに切り離しを指示する。

【０１５１】

【発明の効果】本発明のリング接続を用いたデータ転送方法及び情報処理システムによれば、プロセッサを経由するためのレイテンシを短縮することができ、高性能な情報処理システムを実現することができる。

【０１５２】また、リングが通常動作しなくなる切り替え時にもリングよりモジュールに安定したクロックを供給することできる。

【０１５３】さらに、本発明のリング接続を用いたデータ転送方法及び情報処理システムによれば、情報処理システムのスループットを向上させることができる。

【０１５４】さらに、本発明のリング接続を用いたデータ転送方法及び情報処理システムによれば、リングの構成変更に伴うシステム停止時間を短縮することができる。

【図面の簡単な説明】

【図１】リング、バス及びスイッチの各プロセッサ間接続技術の比較を示す図である。

【図２】リング接続におけるリングのインターフェイス部の構成を示す図である。

【図３】転送のビット幅が１６ビットのリングを示す図である。

【図４】転送のビット幅が６４ビットのリングを示す図である。

【図５】従来技術のリングのフラグ処理を示す図である。

【図６】本発明のリングのフラグ処理を示す図である。

【図７】本発明のリングの送受信を示す図である。

【図８】従来技術のリングのバッファを示す図である。

【図９】従来技術のリングのクロックの乗り換えを示す図である。

【図１０】本発明のリングの送受信のクロックを示す図である。

【図１１】本発明のクロックの位相の進行を示す図である。

【図１２】本発明のクロックの系統を示す図である。

【図１３】従来技術のリングの転送を示す図である。

【図１４】本発明のリングの転送を示す図である。

【図１５】本発明の各モジュールの動作モードを説明するためのパケットの転送例を示す図である。

【図１６】本発明のスキュー補正を示す図である。

【図１７】本発明のリングの種類を示す図である。

【図１８】本発明のブロックパケットリングを示す図である。

【図１９】本発明の並列転送の例を示す図である。

【図２０】本発明の並列転送パケットの選択論理方式を示す図である。

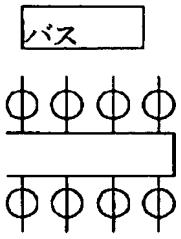
【図２１】本発明のディレクトリリングを示す図である。

【図２２】本発明のディレクトリへのアクセスを示す図

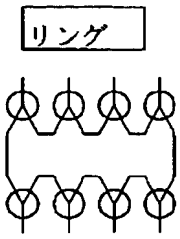
である。
【図２３】従来技術の冗長リングを示す図である。
【図２４】本発明の基本リングと制御リングを示す図である。
【図２５】本発明によるレイテンシの改善内容を示す図である。
【図２６】本発明による動作周波数に改善内容を示す図である。
【図２７】本発明による増設性を示す図である。
【図２８】本発明の基本リングを示す図である。
【図２９】本発明の制御リングを示す。
【図３０】従来の送信権獲得方式を示す。
【図３１】本発明のゲート２段で転送する方式を示す。
【図３２】従来の受信方式を示す図である。
【図３３】本発明の送受信を並行する方式を示す図である。
【図３４】本発明の選択線、アドレス線とデータ線との関係を示す図である。
【図３５】本発明の受信方式を示す図である。
【図３６】本発明の受信動作のタイムチャートを示す図である。
【図３７】本発明の送信動作の概念図を示す図である。
【図３８】本発明の送信動作を示す図である。
【図３９】本発明の送信動作の一例を示す図である。
【図４０】スキュー補正しない例を示す図である。
【図４１】本発明のクロックのスキュー補正する例を示す図である。
【図４２】本発明の受信線毎にスキュー補正する例を示す図である。
【図４３】従来方式の送受信方式を示す図である。
【図４４】本発明のリングのクロックに同期して動作する方式を示す図である。
【図４５】本発明の転送モードの送信動作の概念図を示す図である。
【図４６】本発明のゲートの遅延を利用した位相調整回路を示す図である。
【図４７】本発明のサイクル調整回路を示す図である。
【図４８】本発明の受信線毎にスキュー補正する転送モードの概念図を示す。
【図４９】本発明の送達確認の説明を示す図である。
【図５０】本発明の独立した送信権が２つのブロック転送を示す図である。
【図５１】本発明のリング１周より大きなブロック転送を示す図である。
【図５２】本発明のディレクトリリングのキャッシュがある場合の動作の例を示す図である。
【図５３】本発明のディレクトリリングのキャッシュのない場合の動作の例を示す図である。
【図５４】本発明のディレクトリリングの未確定で再要求の動作例を示す図である。
【図５５】本発明の各モジュールにおいて受信される順序を示す図である。
【図５６】本発明の論理的順序の時刻の決定法を示す図である。
【図５７】本発明の送信権獲得論理を示す図である。
【図５８】本発明の８モジュールの送信権獲得論理を示す図である。
【図５９】本発明の接続数の多い送信権処理を示す図である。
【図６０】本発明の並列転送可能ブロックの選別方式を示す図である。
【図６１】本発明の並列転送ブロックの選択を示す図である。
【図６２】本発明の階層接続の接続方法を示す図である。
【図６３】本発明の階層接続の場合のキャッシュ可能なアクセスをする接続機構を示す図である。
【図６４】本発明の階層接続のアドレスの関係を示す図である。
【図６５】本発明のローカルディレクトリを示す図である。
【図６６】本発明のアドレスの拡張を示す図である。
【図６７】本発明のアドレス変換テーブルを示す図である。
【図６８】本発明のリモートディレクトリを示す図である。
【図６９】本発明におけるディレクトリにより分けられるローカルメモリの種別を示す図である。
【図７０】本発明の双方向リングの弁別回路を示す図である。
【図７１】本発明の双方向リングのモジュール間接続をしめす図である。
【図７２】本発明のモジュールの挿入、引き抜きの早期検出方式を示す図である。
【図７３】本発明のリングインターフェースの具体例を示す図である。
【符号の説明】
１１…データ線、１２、２１、２２…論理積、１３…論理和回路、１４、１５…受信ラッチ、１６…自分アドレス、１７…一致論理、１８、１９、２０…ラッチ、２３…受信パケット、１０１、１０２、１０４、１０５…モジュール、７０１…主リング、７００、７０４、７０６…副リング、７０３、７０７…ローカルメモリ、７０５、７０６…リモートディレクトリ、７０８、７０９…接続機構、７２０…選択線、７２８、７３６…論理和、７２９、７３７…論理積、７３３…受信線。

【図１】

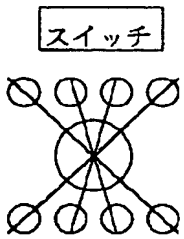
図１



バス



リング



スイッチ

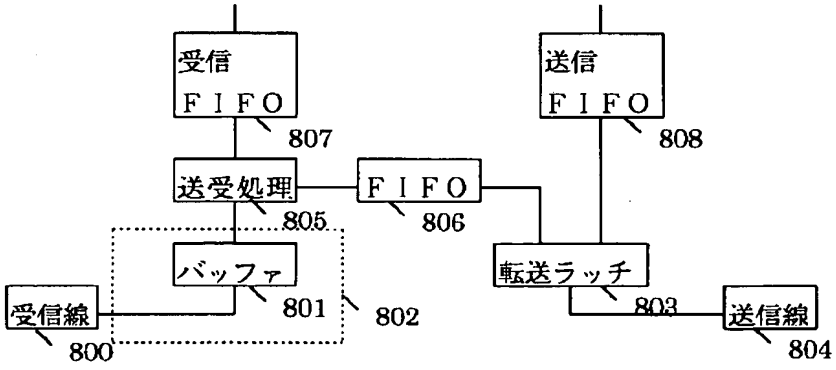
＊は本方式の改善点

○はＬＳＩを示す

小規模向け	中規模向け	大規模向け
・外部回路がない	・外部回路が殆どない	・外部回路が大きい
・全報知の性能高い	・全報知の性能高い	・全報知の性能低い
・回路量少ない	・回路量多い	・回路量中位
・ピン数少ない	＊ピン数は中位	・ピン数少ない
・レイテンシ小さい	＊レイテンシ大きい	・レイテンシ小さい
・動作周波数低い	＊動作周波数高い	・動作周波数中位
・並列転送不可	＊並列転送不可	・完全な並列転送
・構成変更容易	＊構成変更面倒	・構成変更容易

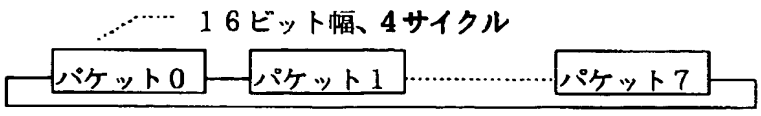
【図２】

図２



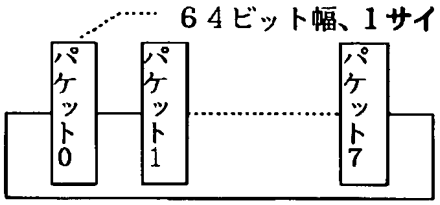
【図３】

図３



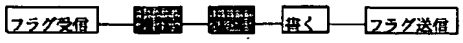
【図４】

図４



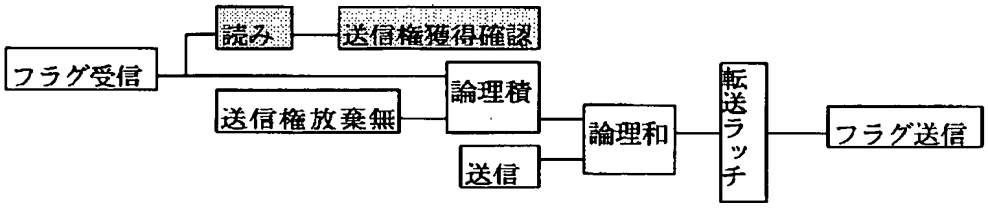
【図５】

図５



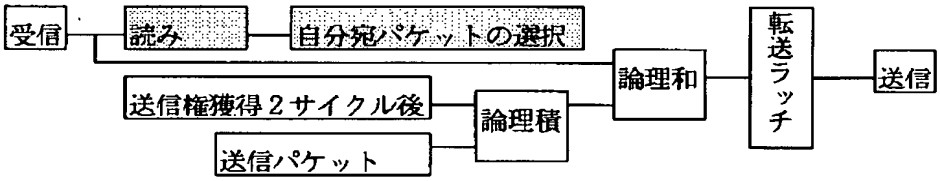
【図６】

図６



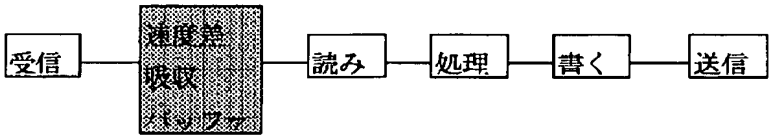
【図７】

図７



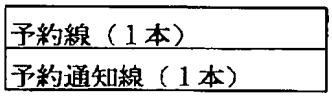
【図８】

図８



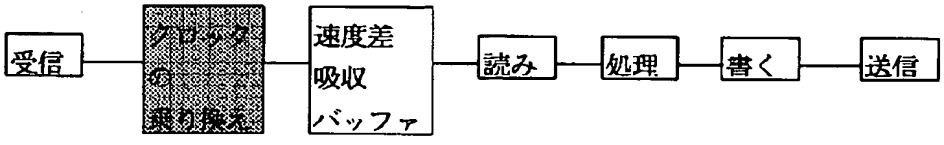
【図１８】

図１８



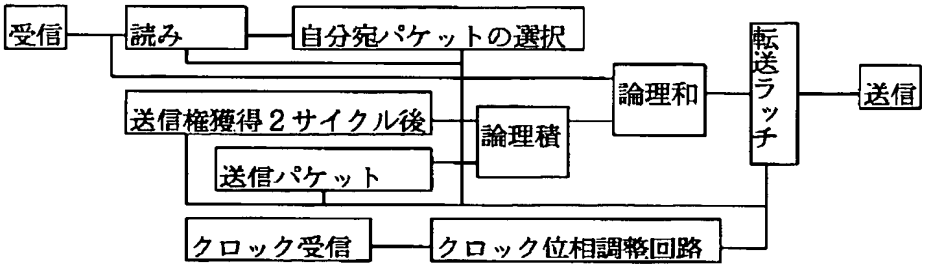
【図９】

図９



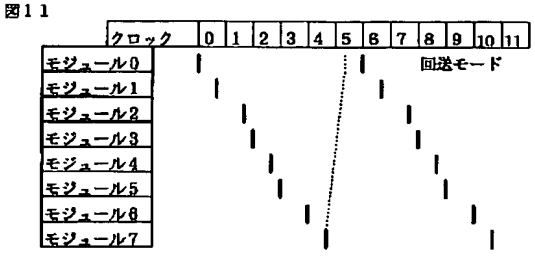
【図１０】

図１０



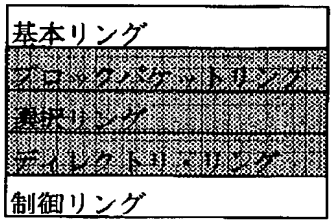
【図１１】

図１１



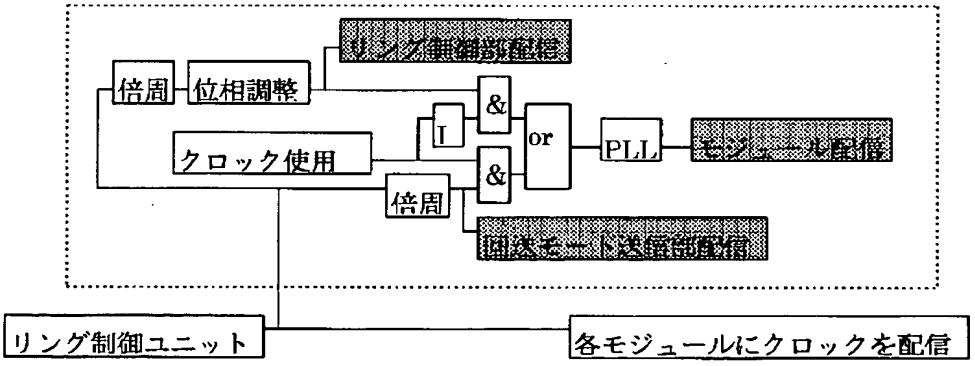
【図１７】

図１７



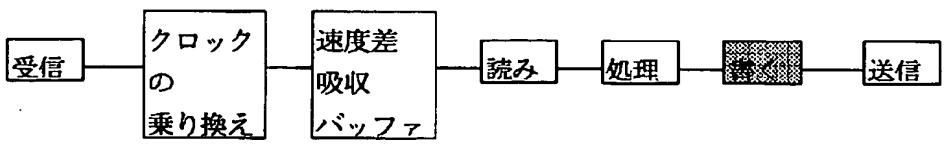
【図１２】

図１２



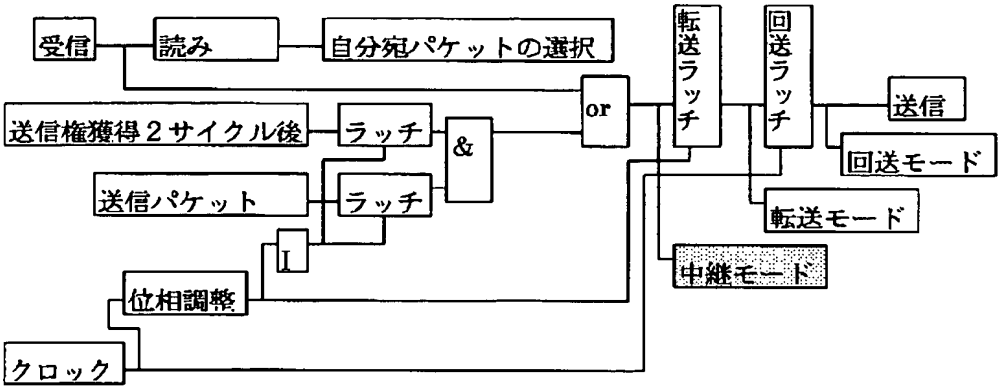
【図１３】

図１３



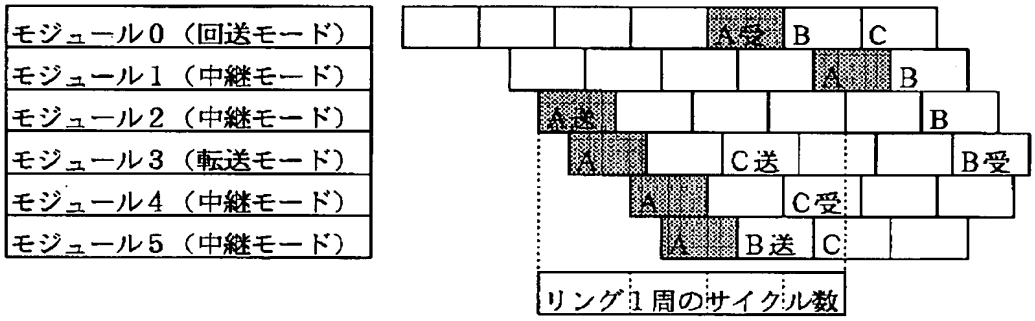
【図１４】

図１４



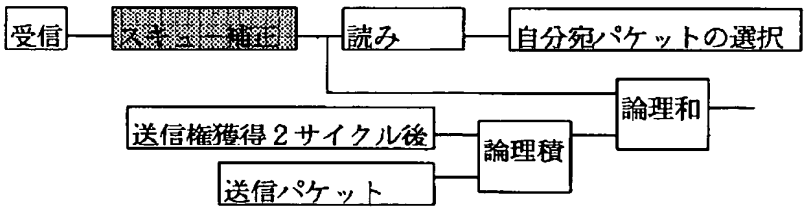
【図１５】

図１５



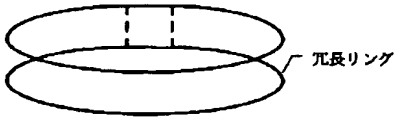
【図１６】

図１６



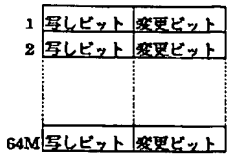
【図２３】

図２３



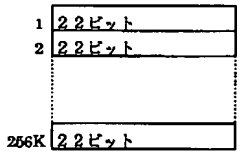
【図６５】

図６５



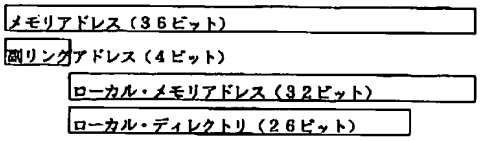
【図６７】

図６７



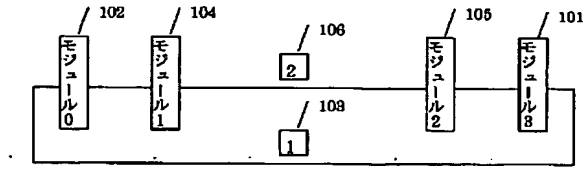
【図６４】

図６４



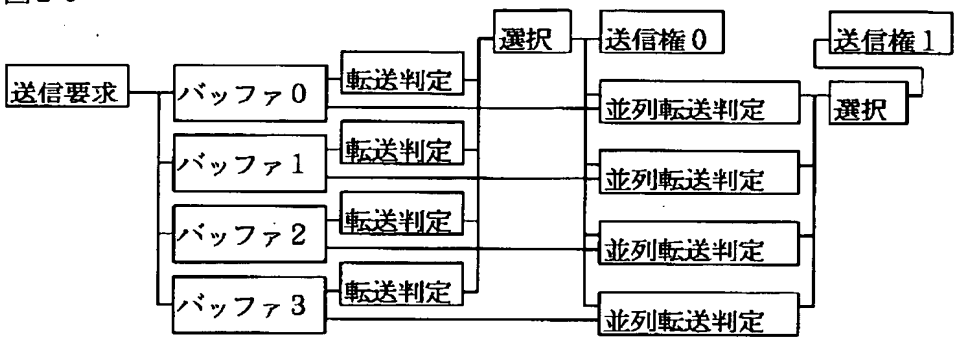
【図１９】

図19



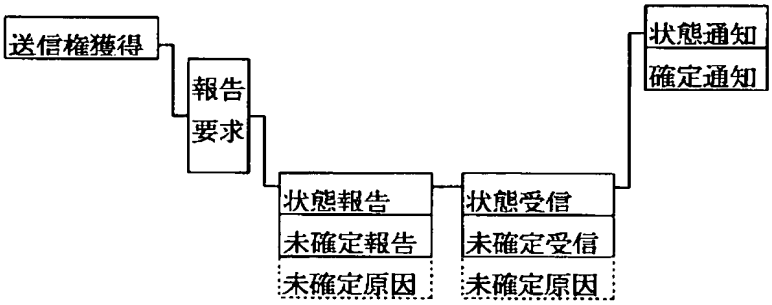
【図２０】

図２０



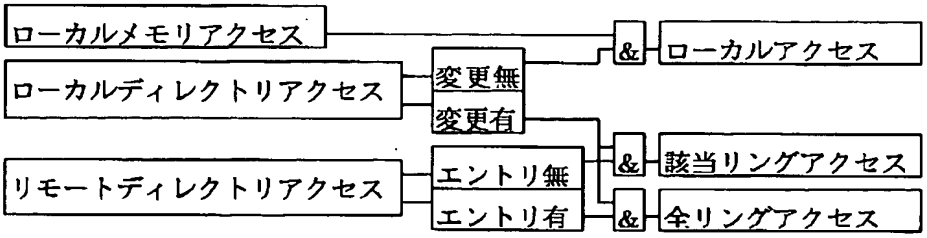
【図２１】

図２１



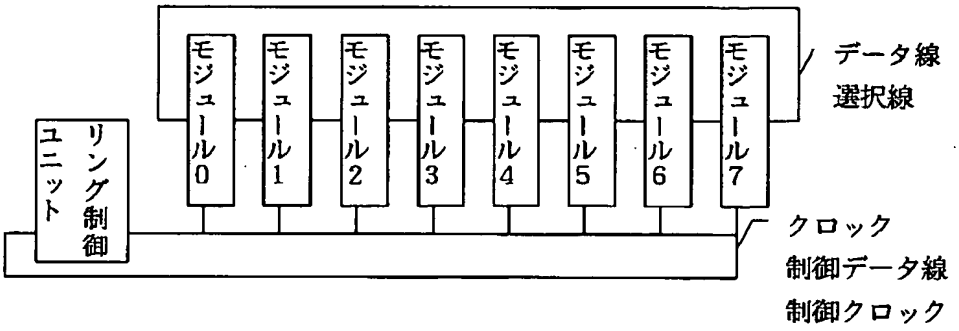
【図２２】

図２２



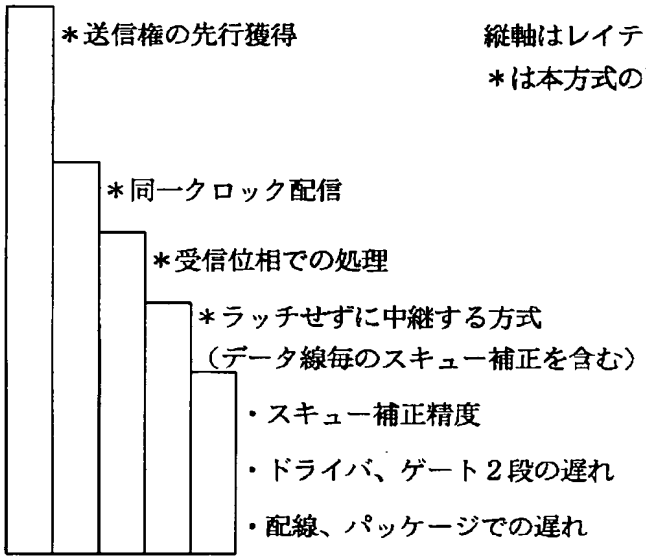
【図２４】

図２４



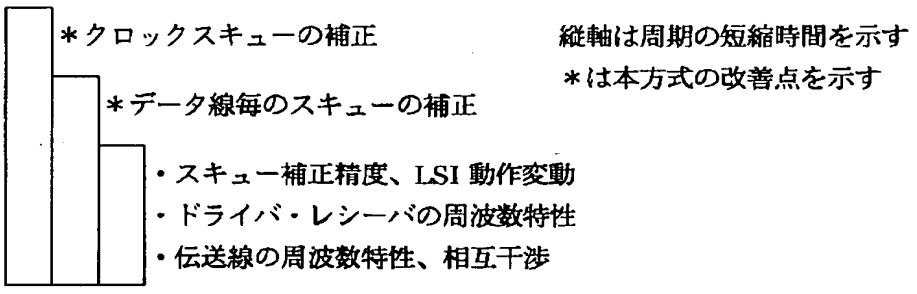
【図２５】

図２５



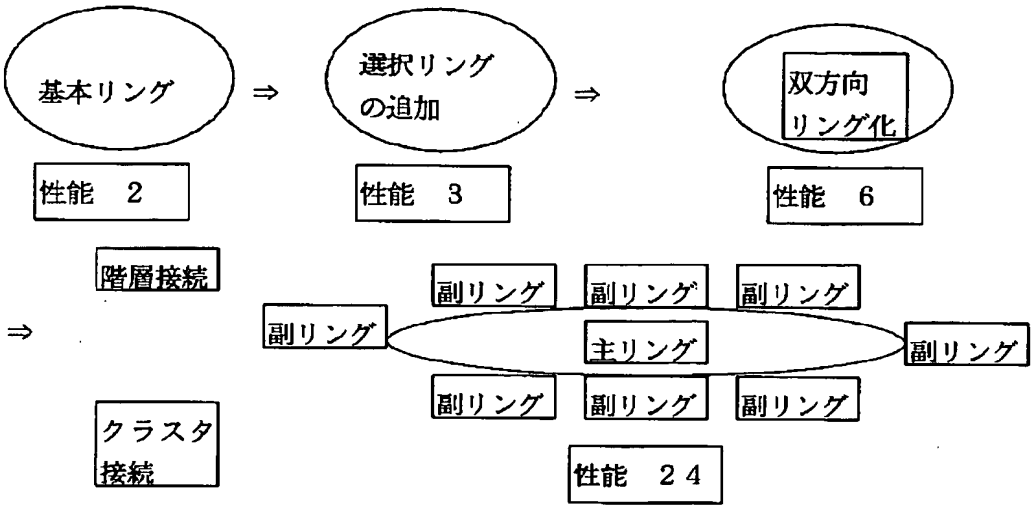
【図２６】

図２６



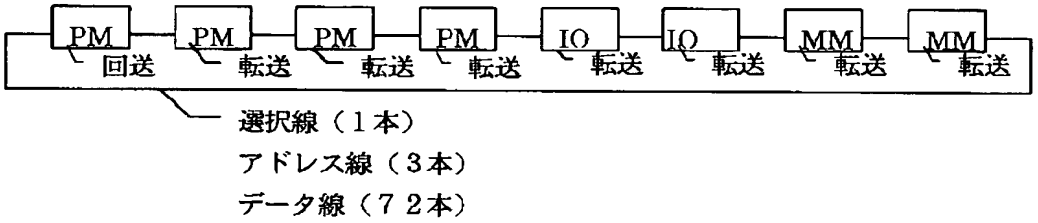
【図２７】

図２７



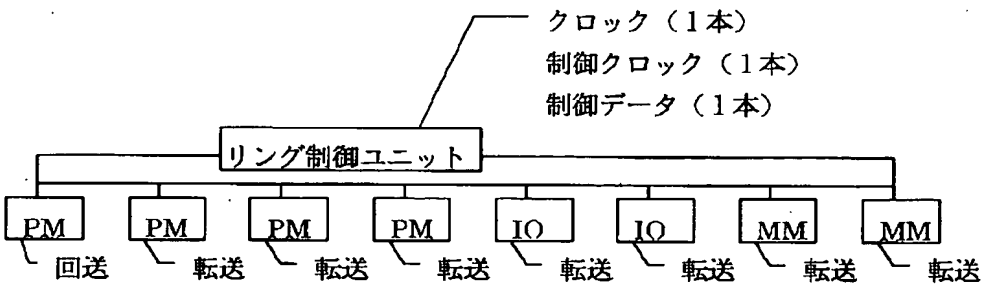
【図２８】

図２８



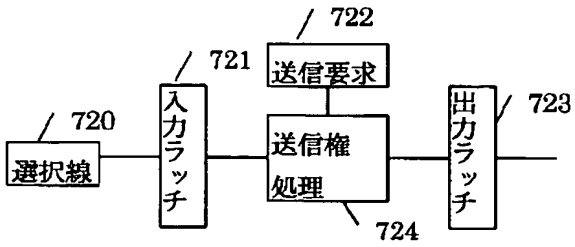
【図２９】

図２９



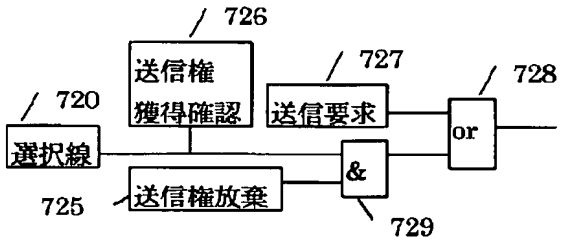
【図３０】

図３０



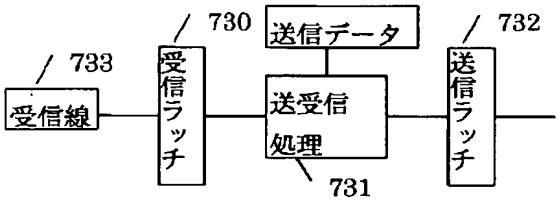
【図３１】

図３１



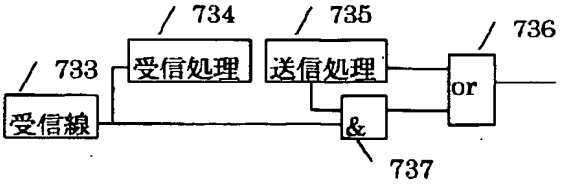
【図３２】

図３２



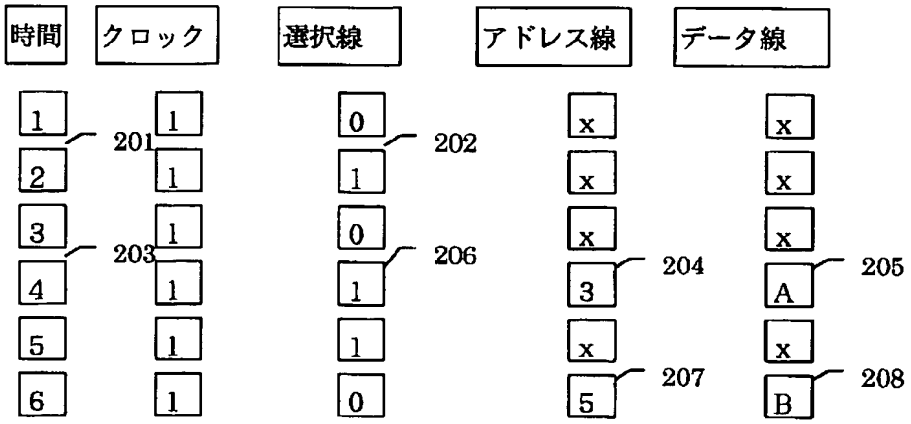
【図３３】

図３３



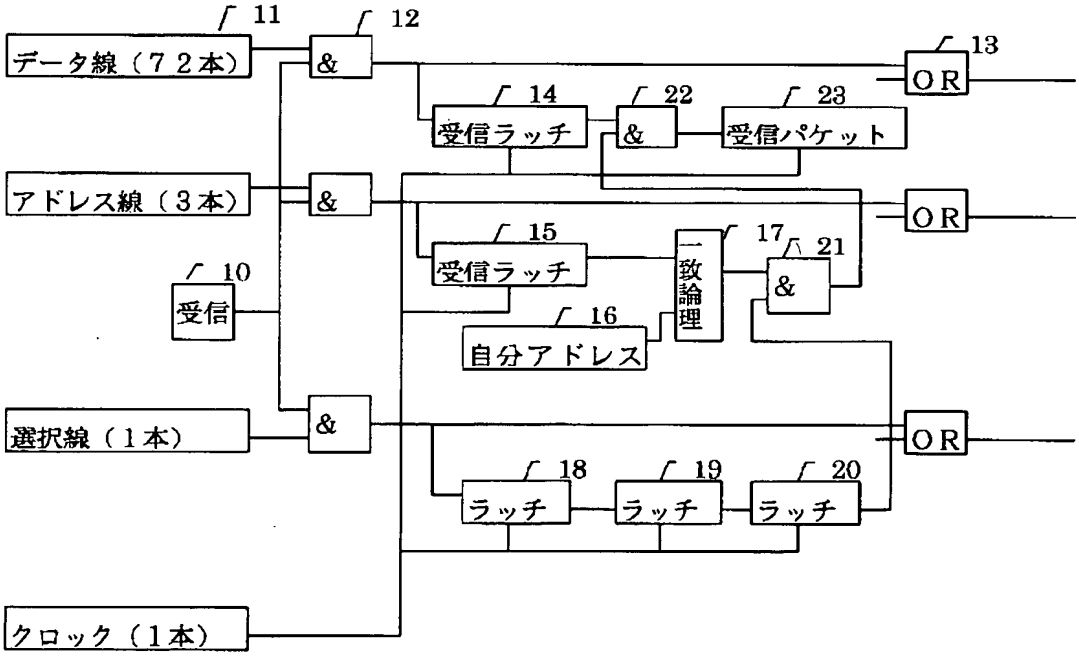
【図３４】

図３４



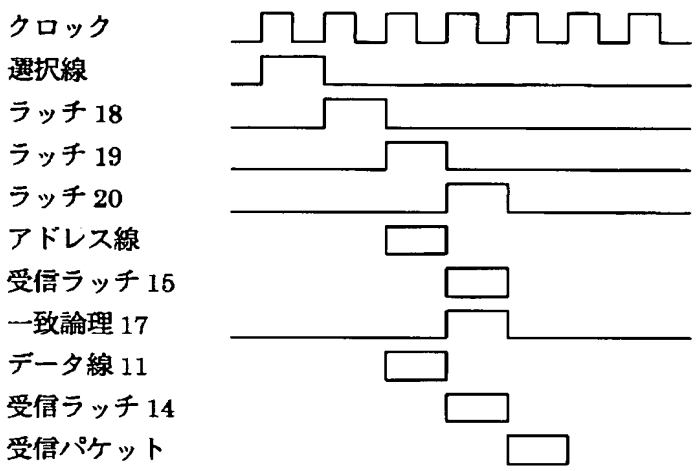
【図３５】

図３５



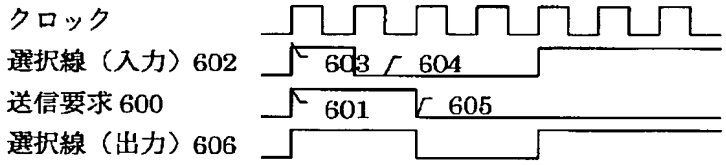
【図３６】

図３６



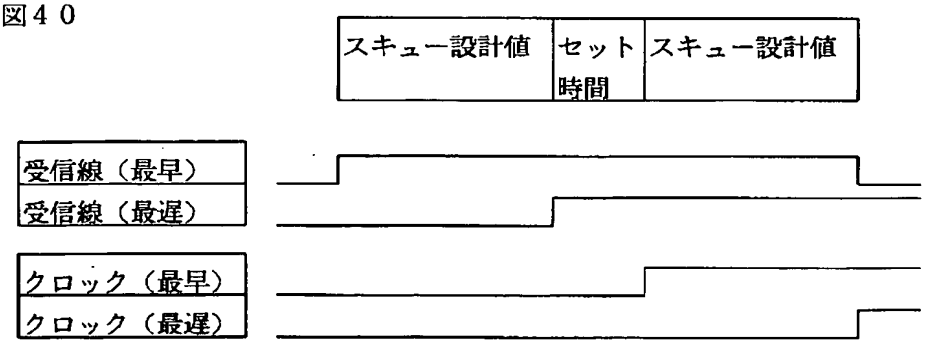
【図３７】

図３７



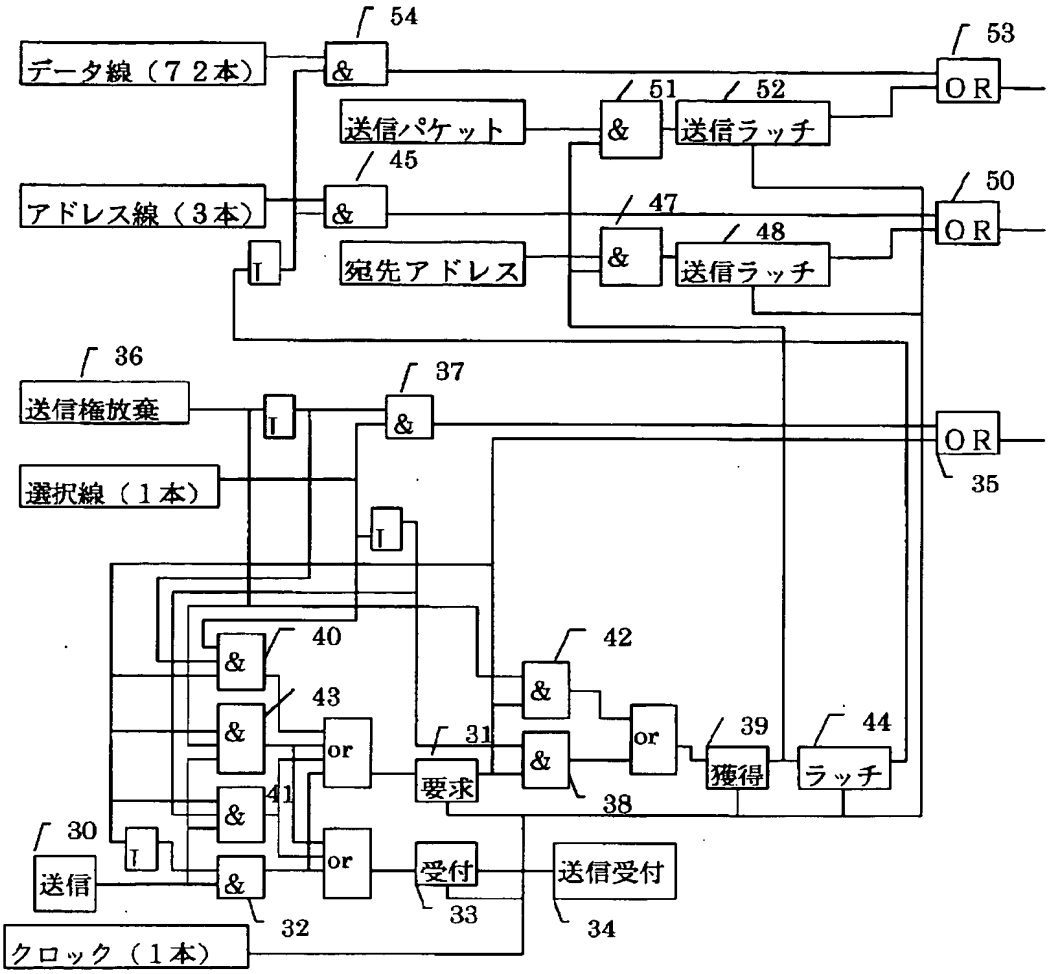
【図４０】

図４０



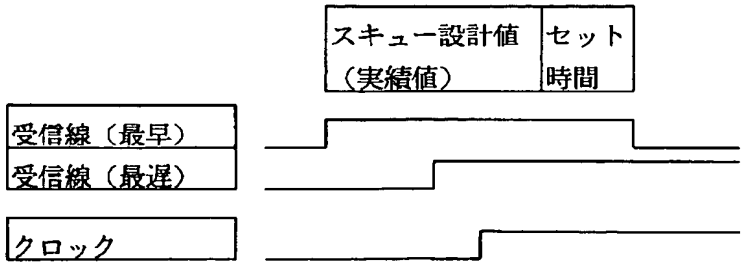
【図３８】

図３８



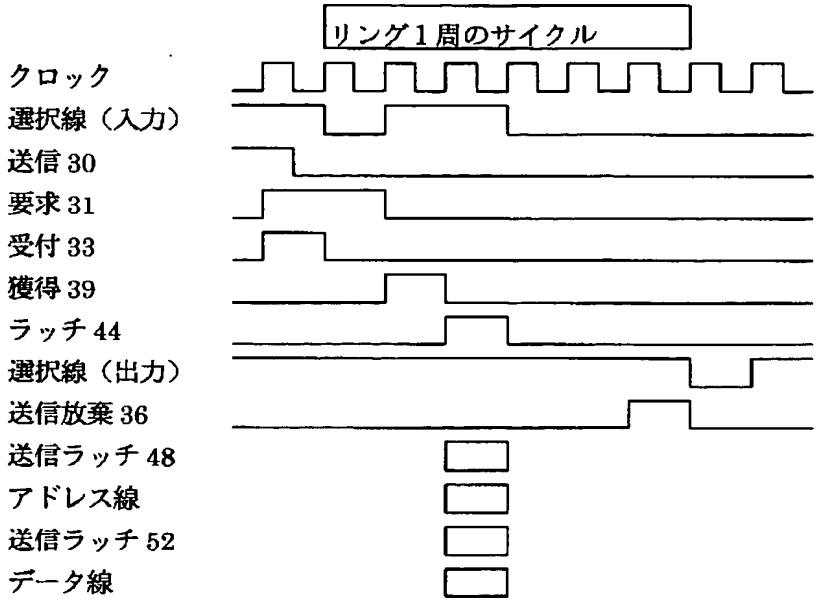
【図４１】

図４１



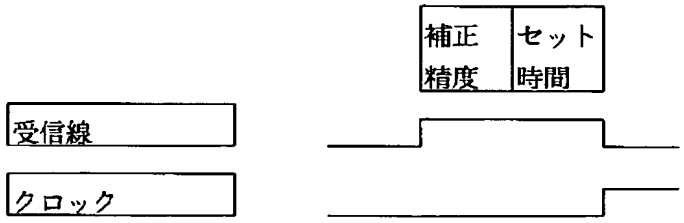
【図３９】

図３９



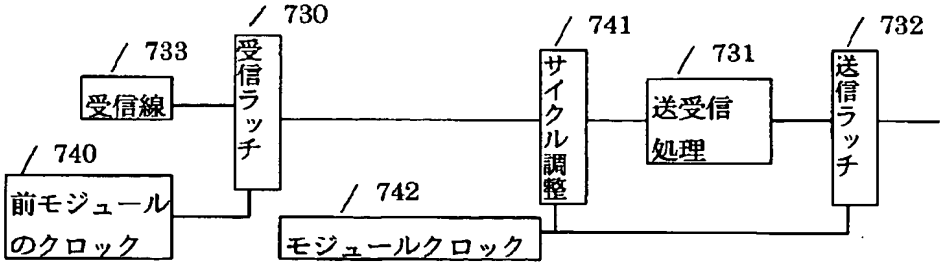
【図４２】

図４２



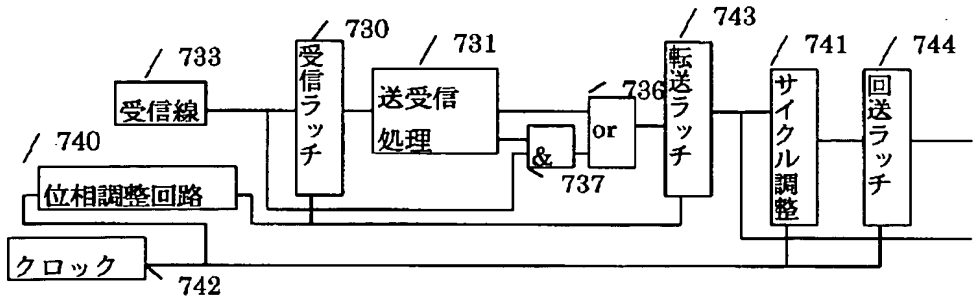
【図４３】

図４３



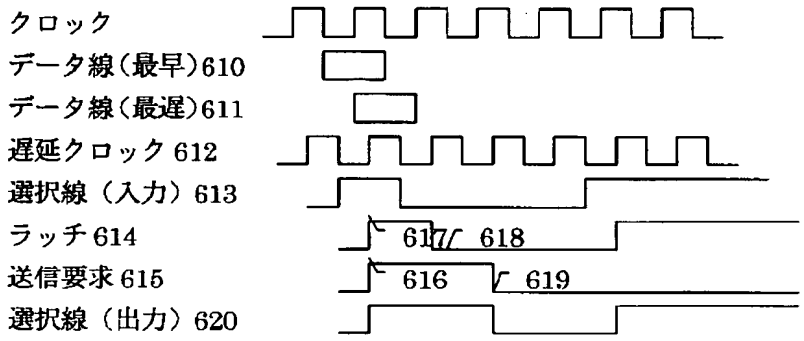
【図４４】

図４４



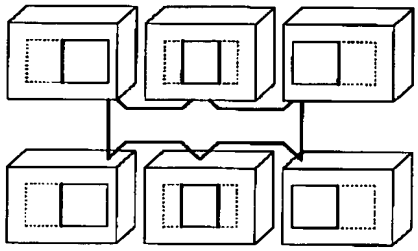
【図４５】

図４５



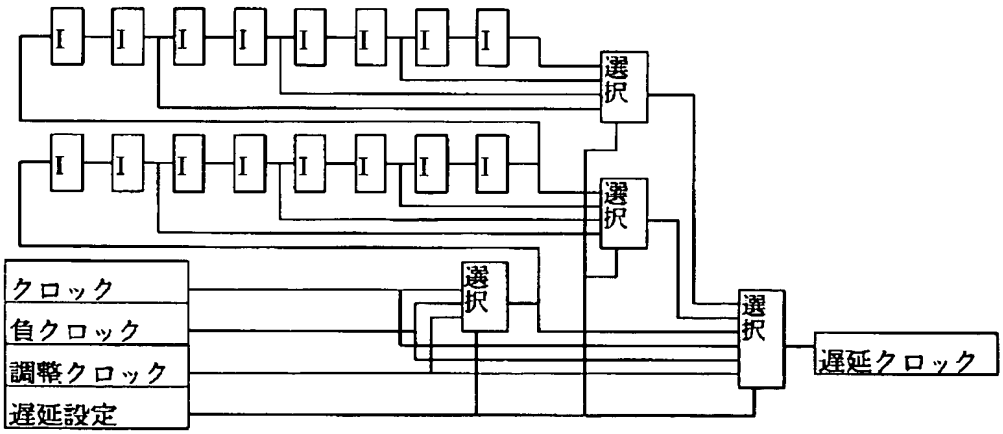
【図６２】

図６２



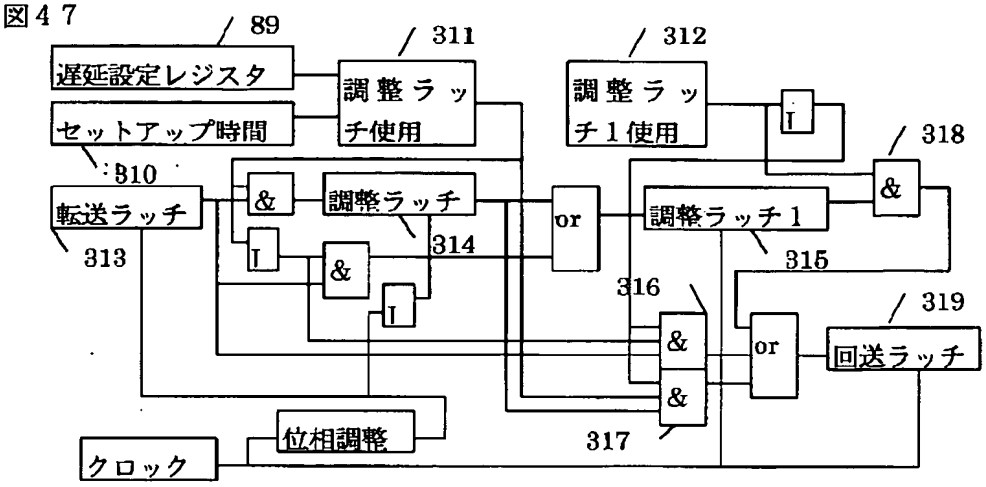
【図４６】

図４６



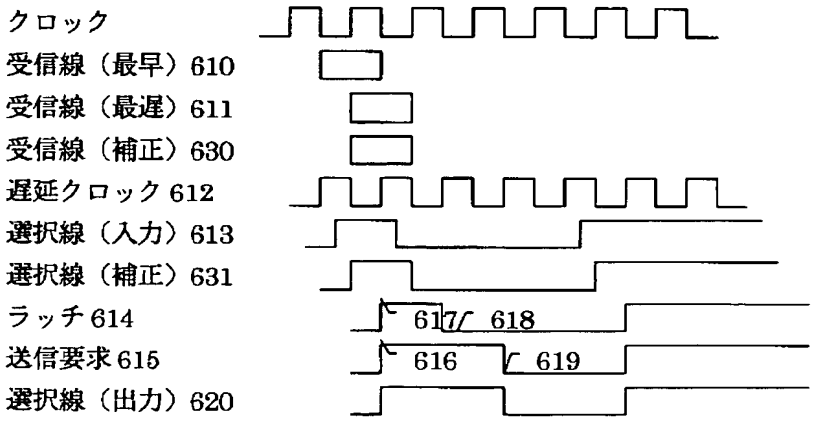
【図４７】

図４７



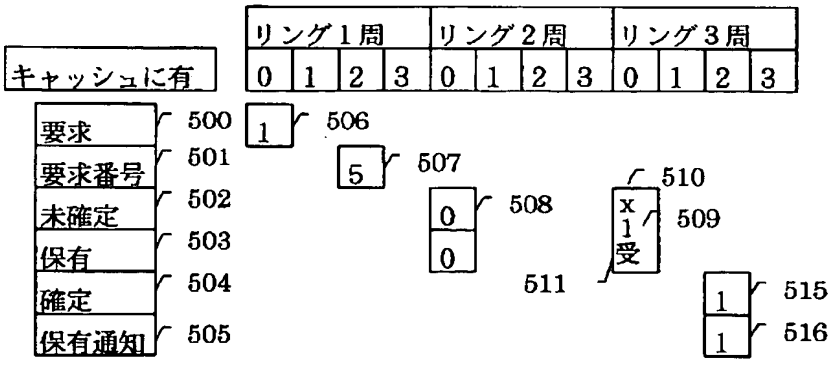
【図４８】

図４８



【図５２】

図５２



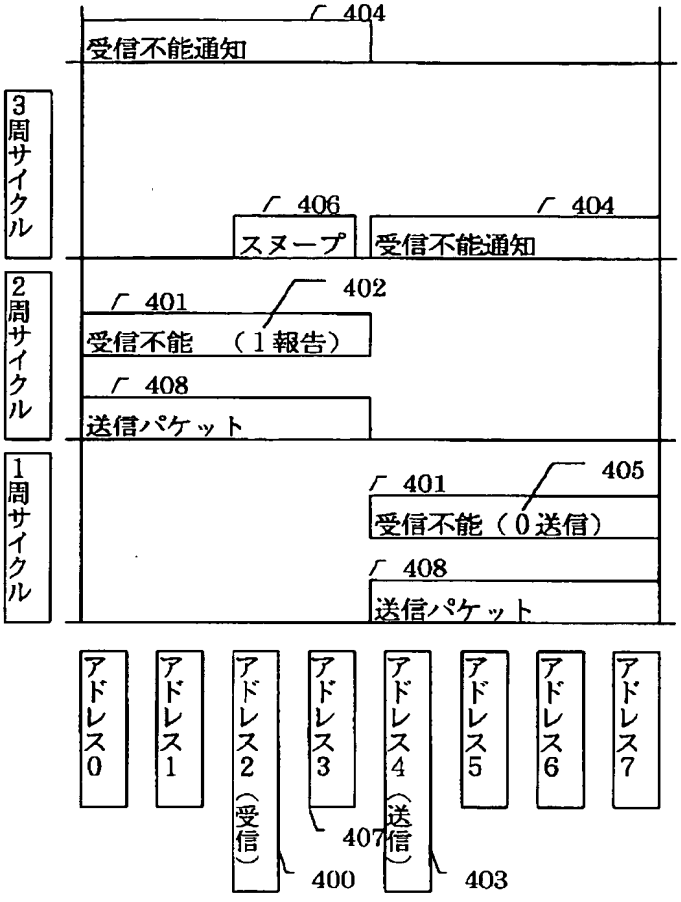
【図６９】

図６９

写し無し（N）	
写し有り（C）	写しベクトル有り（CH）
変更有り（M）	写しベクトル有り（MH）

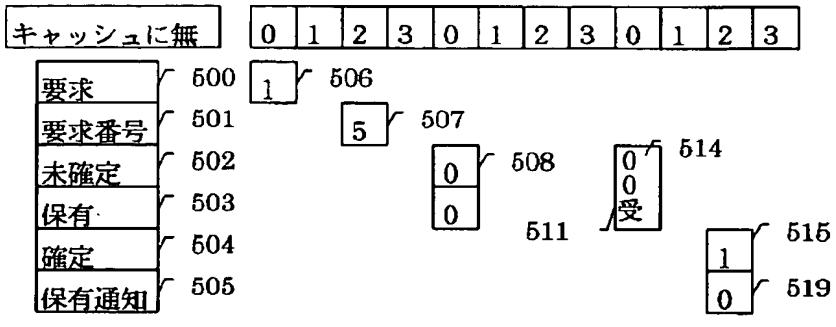
【図４９】

図４９



【図５３】

図５３



【図５０】

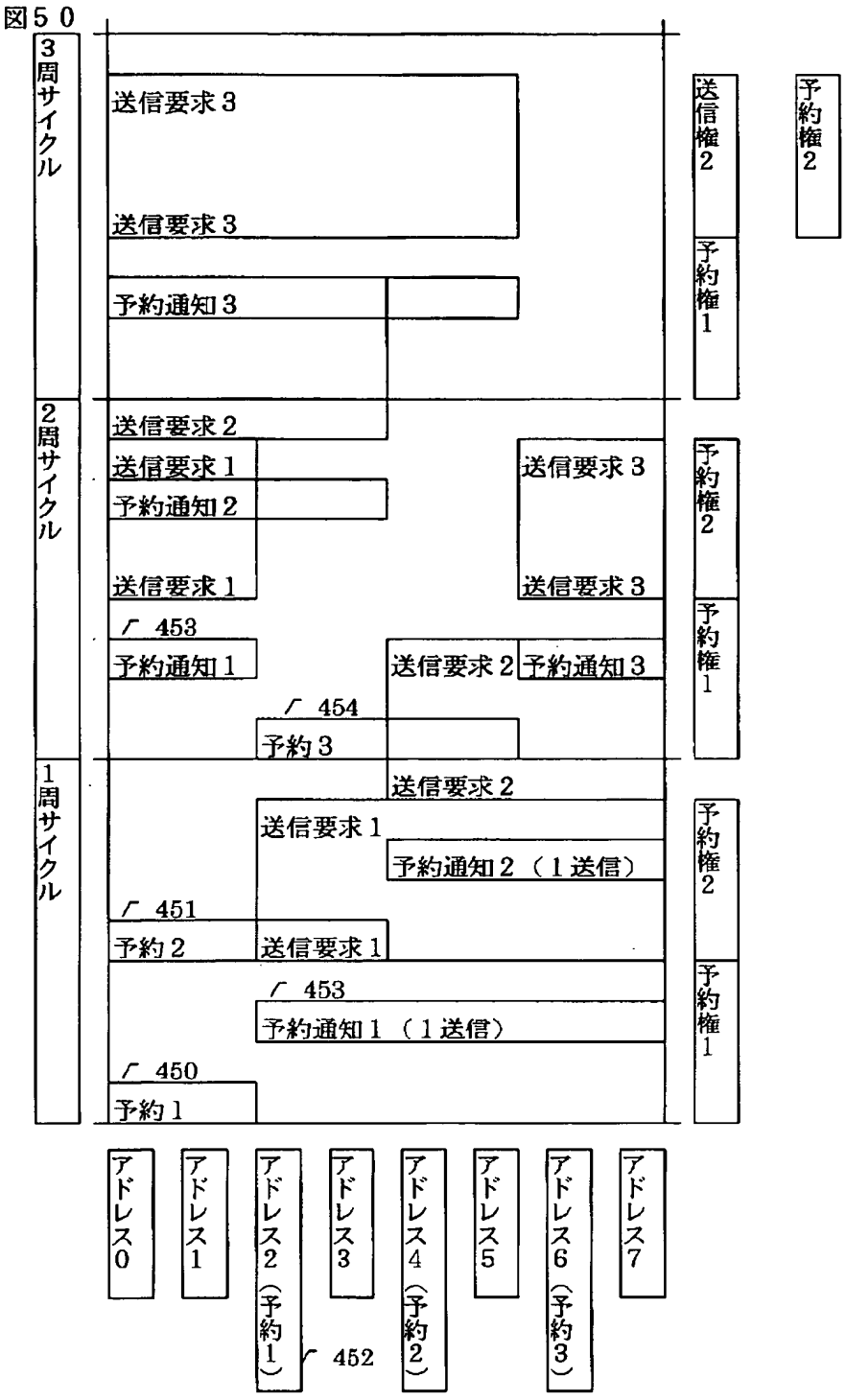
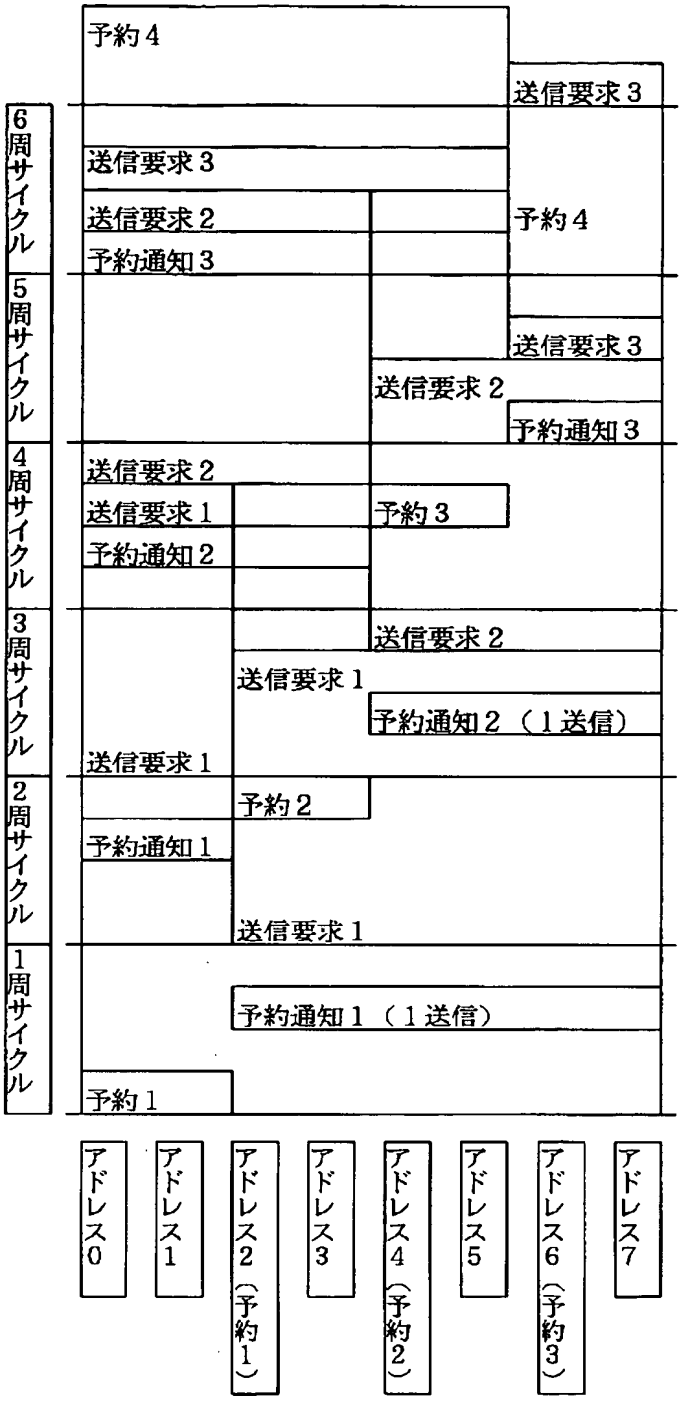


図５０
3周サイクル
2周サイクル
1周サイクル
送信要求３
送信要求３
予約通知３
送信要求２
送信要求１
予約通知２
送信要求１
送信要求３
送信要求３
453
予約通知１
送信要求２
予約通知３
454
予約３
送信要求２
送信要求１
予約通知２（１送信）
451
予約２
送信要求１
453
予約通知１（１送信）
450
予約１
送信権２
予約権１
予約権２
予約権２
予約権１
予約権２
予約権１
アドレス０
アドレス１
アドレス２（予約１）
アドレス３
アドレス４（予約２）
アドレス５
アドレス６（予約３）
アドレス７
452

【図５１】

図５１



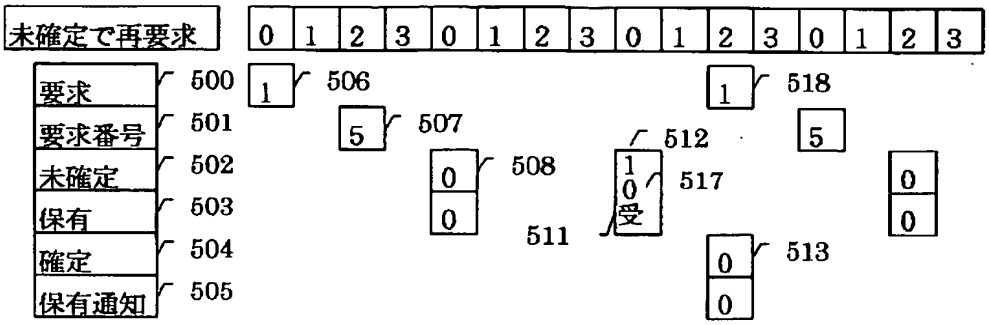
【図６８】

図６８

1	アドレス（４／８ビット）	写しベクトル（８ビット）
2	アドレス（４／８ビット）	写しベクトル（８ビット）
4M	アドレス（４／８ビット）	写しベクトル（８ビット）

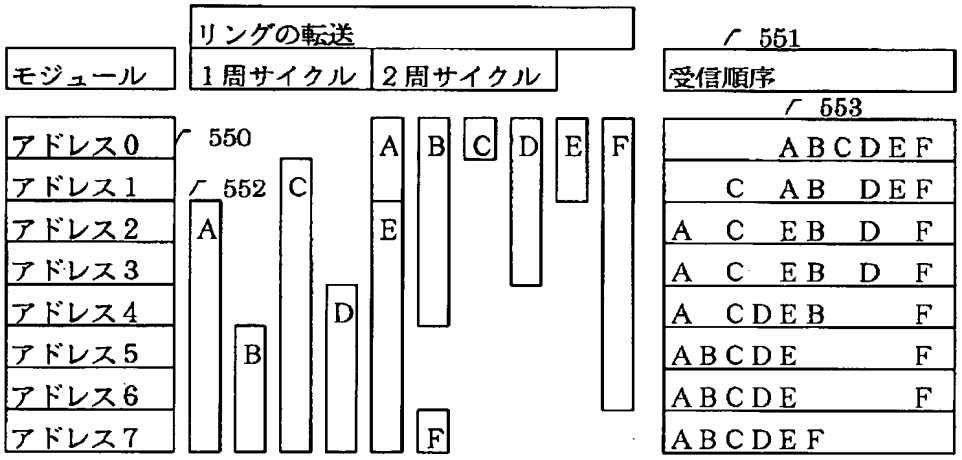
【図５４】

図５４



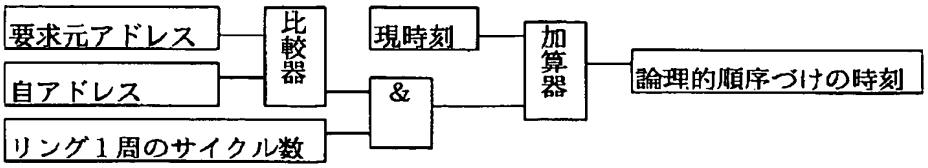
【図５５】

図５５



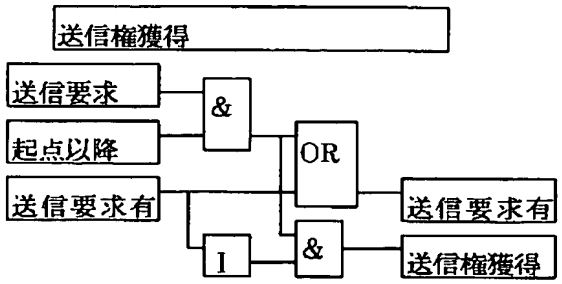
【図５６】

図５６



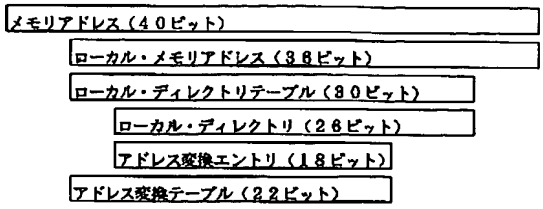
【図５７】

図５７



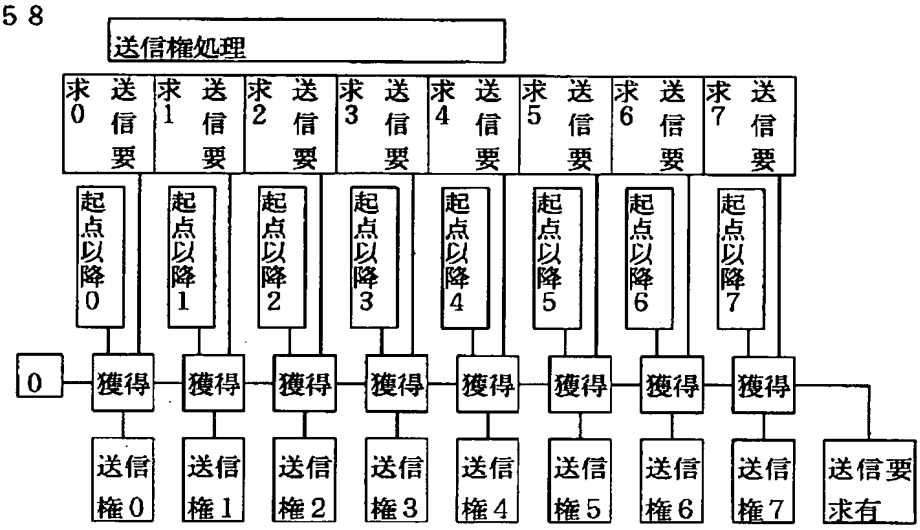
【図６６】

図６６



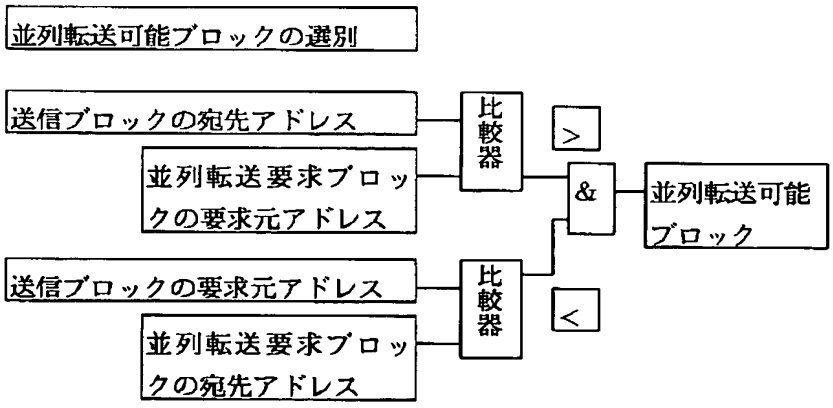
【図５８】

図５８



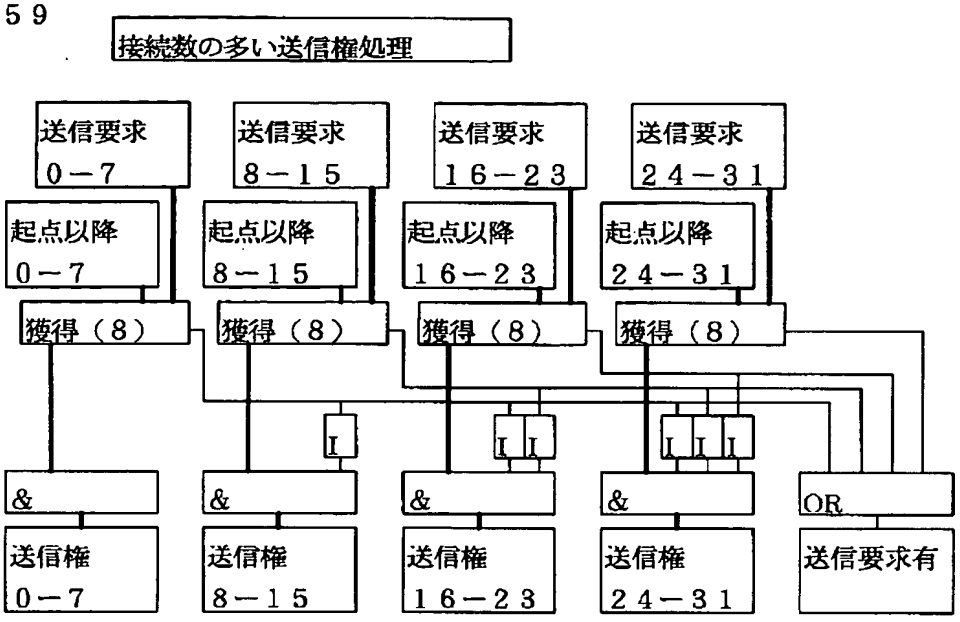
【図６０】

図６０



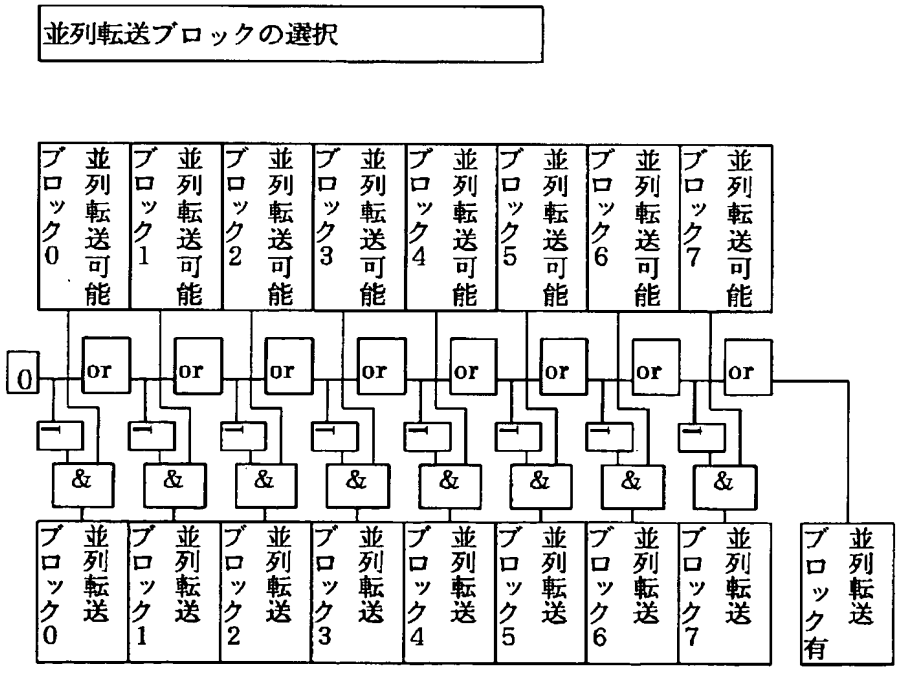
【図５９】

図５９



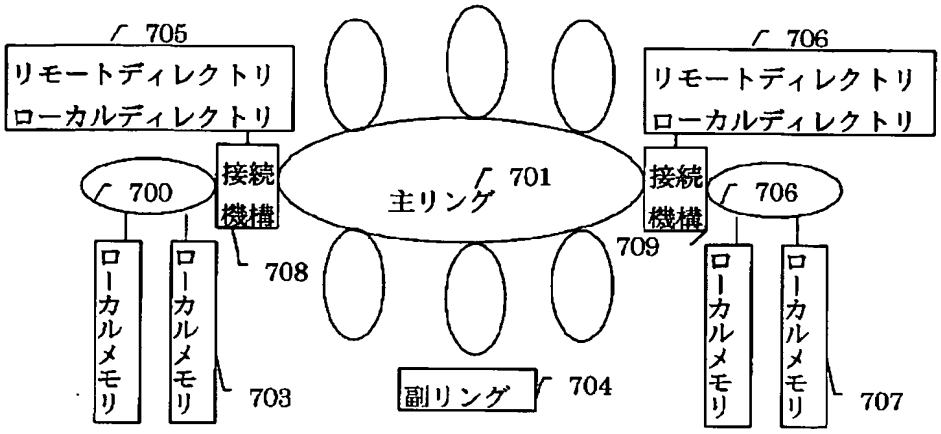
【図６１】

図６１



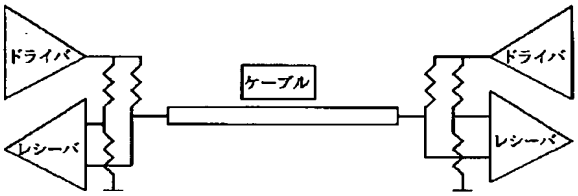
【図６３】

図６３



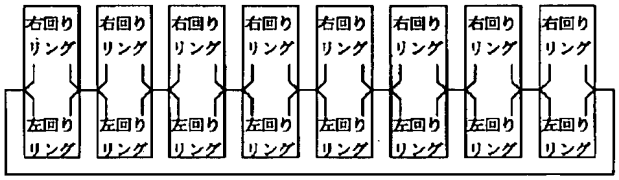
【図７０】

図７０



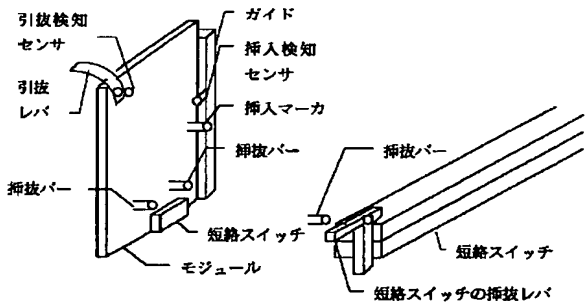
【図７１】

図７１



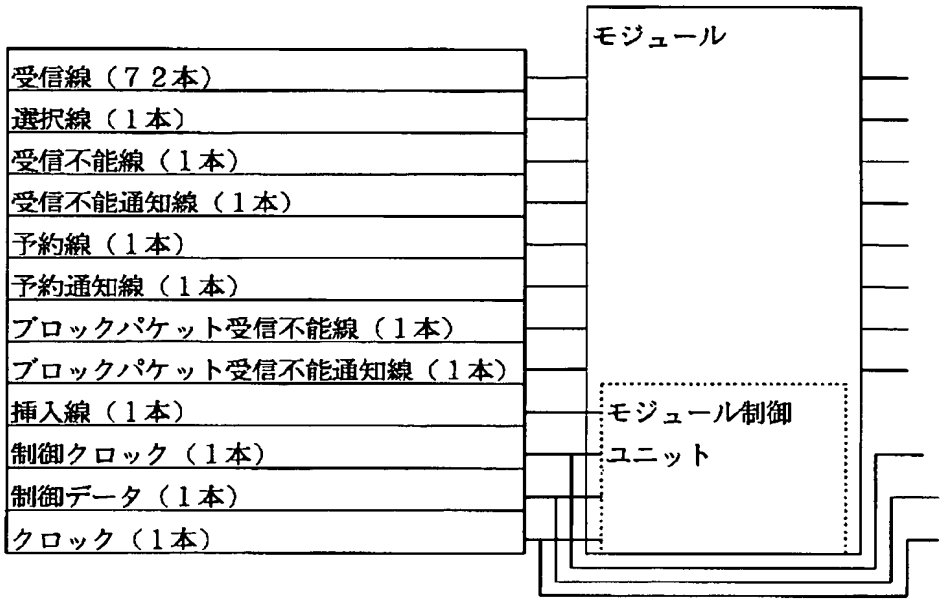
【図７２】

図７２



【図７３】

図７３


モジュール
受信線（７２本）
選択線（１本）
受信不能線（１本）
受信不能通知線（１本）
予約線（１本）
予約通知線（１本）
ブロックパケット受信不能線（１本）
ブロックパケット受信不能通知線（１本）
挿入線（１本）
制御クロック（１本）
制御データ（１本）
クロック（１本）
モジュール制御
ユニット